

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 13, 2006, 09:34:18 ; Search time 189.5 Seconds
(without alignments)

4953.591 Million cell updates/sec

Title: US-10-701-844-1

Perfect score: 7883

Sequence: 1 999caaaactctccccccg.....gcactcctaagaaagaattc 4435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DRV=xlp
-O=/abes/ABSSWEB spool/US10701844/runat 12052006 165420 26277/app query.fasta_1
-DB=UniProt -QEMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p
-USER=US10701844 @CGN 1.1 466 @runat 12052006 165420 26277 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONLQAG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5131.5	65.1	1013	1 PMPG CHLTR	O84879 chlamydia t
2	3846	48.8	987	1 PMPG CHLMU	O9p145 chlamydia m
3	1860.5	23.6	1011	2 Q823X5 CHLCV	O823x5 chlamydropi
4	1777	22.5	1024	2 Q5L6J3 CHLAB	Q5L6j3 chlamydropi
5	1589	20.2	1016	1 PMPH CHLTR	O84880 chlamydia t
6	1566	19.9	981	2 Q84F07 CHLTR	O84fu7 chlamydia t
7	1551.5	19.7	980	2 Q84F06 CHLTR	O84fu6 chlamydia t
8	1551.5	19.7	980	2 Q84F08 CHLTR	O84fu8 chlamydia t
9	1459	18.5	989	2 Q83U76 CHLTR	O83u76 chlamydia t
10	1459	18.5	989	2 Q84F09 CHLTR	O84fu9 chlamydia t
11	1452	18.4	991	2 Q83UW2 CHLTR	O83uw2 chlamydia t
12	1452	18.4	991	2 Q84FV0 CHLTR	O84fv0 chlamydia t
13	1452	18.4	991	2 Q84FV1 CHLTR	O84fv1 chlamydia t
14	1404.5	17.8	993	2 Q83TJ6 CHLTR	O83tj6 chlamydia t
15	1404.5	17.8	993	2 Q84FV2 CHLTR	O84fv2 chlamydia t
16	1365.5	17.3	973	1 PMP13 CHLPN	Q92896 chlamydia p

17	1174	14.9	980	1 PMPH CHLMU	Q9p144 chlamydia m
18	1158	14.7	942	2 Q823X1 CHLCV	Q823x1 chlamydropi
19	1122	14.2	926	2 P71135 CHLAB	P71135 chlamydropi
20	1122	14.2	926	2 Q5L6J3 CHLAB	Q5L6j3 chlamydropi
21	1101	14.0	928	1 PMP10 CHLPN	Q92b65 chlamydia p
22	1099	13.9	926	2 Q823W9 CHLCV	O823w9 chlamydropi
23	1090	13.8	928	1 PMP11 CHLPN	O86164 chlamydia p
24	1079	13.7	930	1 PMP8 CHLPN	Q92393 chlamydia p
25	1077	13.7	1378	2 Q5L6J2 CHLAB	Q5L6j2 chlamydropi
26	1040	13.2	936	1 PMP7 CHLPN	Q92898 chlamydia p
27	1021	13.0	928	1 PMP9 CHLPN	Q92398 chlamydia p
28	1015	12.9	1276	1 PMP6 CHLPN	Q92899 chlamydia p
29	1008	12.8	866	2 Q823X0 CHLCV	Q823x0 chlamydropi
30	979.5	12.4	772	2 Q9RB71 CHLPN	Q92b71 chlamydia p
31	973	12.3	922	1 PMP1 CHLPN	Q92895 chlamydropi
32	971.5	12.3	841	2 Q822Q5 CHLCV	Q822q5 chlamydropi
33	950	12.1	847	2 P71132 CHLAB	P71132 chlamydropi
34	949.5	12.0	839	2 P77792 CHLAB	P77792 chlamydropi
35	939.5	11.9	841	1 PMP2 CHLPN	Q923a1 chlamydia p
36	897	11.4	846	2 P71133 CHLAB	P71133 chlamydropi
37	803	10.2	843	2 Q823X2 CHLCV	Q823x2 chlamydropi
38	789	10.0	867	1 PMP1 CHLMU	Q9p141 chlamydia m
39	787.5	10.0	602	2 Q8VU49 CHLPS	Q8vu49 chlamydia p
40	785.5	10.0	843	2 Q823X3 CHLCV	Q823x3 chlamydropi
41	785	10.0	849	2 Q5L6J5 CHLAB	Q5L6j5 chlamydropi
42	784.5	10.0	878	2 Q4VWS6 CHLTR	Q4vws6 chlamydia t
43	781.5	9.9	878	2 Q4VWR2 CHLTR	Q4vwr2 chlamydia t
44	779.5	9.9	878	2 Q4VWS3 CHLTR	Q4vws3 chlamydia t
45	779.5	9.9	878	2 Q4VWS8 CHLTR	Q4vws8 chlamydia t

ALIGNMENTS

RESULT 1

PMPG_CHLTR

ID PMPG CHLTR STANDARD; PRT; 1013 AA.
AC O84879;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G).
DE Name=pmpG; OrderedLocusNames=CT871;
GN Chlamydia trachomatis.
OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OC NCBI_taxid=813;
RX STRAIN=D/UN-3/Cx;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC MEDLINE=99000809; PubMed=97841136; DOI=10.1126/science.282.5389.754;
RX Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;
RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";
RT Science 282:754-759(1998).
RL -|- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (potential).
CC -|- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -|- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC EMBL; AB001360; AAC68469.1; -; Genomic_DNA.
DR PIR; G71460; G71460.
DR PHCI-2DPAGE; O84879; -
DR InterPro; IPR005546; Auto_transpbtbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP_M.

DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 9.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 5.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 27 Potential.
FT CHAIN 28 1013 Probable outer membrane protein pmpG.
SQ SEQUENCE 1013 AA; 107367 MW; F0927743C0A651DD CRC64;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 1 Gaps: 1

US-10-701-844-1 (1-4435) x PMPG_CHL1TR (1-1013)

QY 382 ATGCAAGCTCTTCCATAAGTCTCTTCTCAATGATCTAGCTTATTTGCTGCTCT 441
DB 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysCysSer 20

QY 442 TTAATGGGGGGGATGCGAGCAAAATCATGTTCTCTCAAGGAATTCAGATGGGAG 501
DB 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40

QY 502 ACGTAACTGATCATTTCCCTATCTCTATAGGAGATCCGAGTGGGACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60

QY 562 TCTGAGGAGAGTAACTTAAATAATCTTGACAATCTTATGACAGCTTGTGCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80

QY 622 TGTTTTGGAACTTATTAGGGAGTTTACTGTTTGGAGAGGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100

QY 682 GAGACATACGACTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGTGATGGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120

QY 742 TTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTACTTGCCTGA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140

QY 802 CTGCTGCTGCAACCACTAATAAGGTAGCCAGACTCCGAGCAACAATCTACACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160

QY 862 AATGCTACTATTATTCTTAAACAGACTTTTGTGTTACTCAATATAGAGAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180

QY 922 TATAGTAATTAGTCTCTGGAGATGGGCGAGCTATAGATGCTTAAAGAGCTTAAACGGTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAlaLeuAspAlaLysSerLeuThrValGln 200

QY 982 GGAATTAGCAAGCTTTGTGTTCTCAAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220

QY 1042 CNAGTAGTCACCAAGTTCTGCTATGCTTAACGAGGCTCTTATGTCCTTTGTAGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240

QY 1102 GTTGACGAGTAAAGGGGGAGGATGCTGCTGTTTCAGGATGGGACAGGAGGTGCA 1161
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260

QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACTCGCGTAGAGTTT 1221
DB 261 SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280

QY 1222 GATCGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACCGGAACGTTGCTTTCTCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300

QY 1282 AATAATGGAACAACTTTGTTTCTCAACATGTTGCTTCTCTGTTTACATTCGCTGAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320

QY 1342 CAACCAACAGTGCAGCAGGCTTCTAATACGAGTAATTAATACCGAGATGCGGAGCTATC 1401
DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340

QY 1402 TTCTGTAAGAAATGTCGCGCAA---GCAGGATCCCAATTAACCTCGATCATGTTCTCTTCAT 1458
DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360

QY 1459 GGAGAGGAGTAGTTTCTTTTAGTACAAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380

QY 1519 GCCAAAAAGCTCTCGGTTGCTTAACCTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578
DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400

QY 1579 GATGCTGAGCGATTTATTAGGAGAATCTGGAGAGCTCGAGTCTAGTTTATCTGCTGATTATGGA 1638
DB 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGluLeuSerLeuSerAlaAspTyrGly 420

QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACACCCAAAGAGAAATGCTGCCGATGTTAAT 1698
DB 421 AspileIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440

QY 1699 GGCCTAATCTGCTCTCACAGCCATTTCCGATGGGATCGGGAGGAGGAAAAATAACGACATTA 1758
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460

QY 1759 AGACCTAAGCAGGCGCATCAGATTCCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480

QY 1819 AACCCAGCAGCGCAGTCTTCCAACTTCTAAAAATTAACGATGCTGAAGGATACACAGGG 1878
DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500

QY 1879 GATATTGTTTGTCTAATGAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938
DB 501 AspileValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520

QY 1939 AGGATTGTTCTTCGTGAAAAGGCAAAATTAATCAGTGAATTTCTTAAGTCACAGAGTGGG 1998
DB 521 ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540

QY 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCAACCAACAG 2058
DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560

QY 2059 CCTCTCGCGCTAATCAGTTGATCACGCTTTCGAATCTGCAATTTGCTCTTTCTCTTTG 2118
DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580

QY 2119 TTAGCAACAATGAGTTAGGAATCTCCTACCAATCTCCAGCGCAAGATTCCTCATCCT 2178
DB 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600

QY 2179 GCAGTCAATTGTCACAACTGCTGGTTCTGTTACAATTAGTGGGCCCTATCTTTTTCAG 2238
DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620

QY 2239 GATTGGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298

Db	621	AspLeuAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnIysIleAsp	640		981	SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer	1000
Qy	2299	GTCTCTGAATTAACAGTTAGGAGTAAGCCCGAGCTAATAGCCCGATCAGATTGACTCTA	2358		3379	CGAGCTATGTTTTCAGTGACGAGTAAGTAGAGTCCGGTTC	3417
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660		1001	ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe	1013
Qy	2359	GGGAATGAGATCGCTTAAGTATGGCTATCAAGNAGCTGGAAGCTGCGTGGATCCTAAT	2418		RESULT 2		
Db	661	GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn	680		PMPG_CHLMU STANDARD; PRT; 987 AA.		
Qy	2419	ACAGCAAAATAAGTCTTCTATCTCTGAAGAGCTACATGAGTAAAGCTGGTATATCTCT	2478		AC Q9PLA5; 16-OCT-2001 (Rel. 40, Created)		
Db	681	ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro	700		DT 16-OCT-2001 (Rel. 40, Last sequence update)		
Qy	2479	GGGCTCTGAGCGAGTACCTCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538		DT 13-SEP-2005 (Rel. 48, Last annotation update)		
Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle	720		DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G).		
Qy	2539	CGATCTGCGGATTCAGCAATTCAGCAAGTGTGGATGGGGCTCTTATTGTCGAGGATTA	2598		GN Name=pmpG; OrderedLocusNames=TC0263;		
Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740		OS Chlamydia muridarum.		
Qy	2599	TGGTCTTCTGAGTTCGAATTTCTCTATCATGACCGGATGCTTTAGTTCAGGATAT	2658		OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
Db	741	TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760		OX NCBI_TaxID=83560;		
Qy	2659	CGGTATATTAGTGGGGTTATTCTCTAGGAGCAAACTCTCTACTTGGATCATCGATGTTT	2718		RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
Db	761	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780		RC STRAIN=MoPr / Ni99;		
Qy	2719	GGTCTAGCATTTACCGAAGTATTGTTAGTATCTAAGATTTATGCTGCTCTCCAT	2778		RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;		
Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn	800		RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg T.R., Berry K.J.,		
Qy	2779	CATCATCTGTCATAGGATCGTTTATCTATCTACCCCAACAGCTTTATGTCATCTAT	2838		RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Kouri H.M., Bowman C.,		
Db	801	HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuGlyGlySerTyr	820		RA Bass S., Linher K.D., Weidman J.F., Kouri H.M., Craven B., Bowman C.,		
Qy	2839	TTGTTCTGGAGATCGTTTATCCGCTAGCTAGCTACGGGTTTGGGAATCAGCATATCAAAACC	2898		RA Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,		
Db	821	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840		RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;		
Qy	2899	TCATATACATTTGCAGAGGAGCGATGTTCTGTTGGGATTAATCTCTCGCTGGAGAG	2958		RT "Genome sequences of Chlamydia trachomatis MoPr and Chlamydia pneumoniae AR39.";		
Db	841	SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu	860		RL Nucleic Acids Res. 28:1397-1406(2000).		
Qy	2959	ATTGGAGCGGATACCGATTTGATTTACTCTCACTAGCTCTATTTTGAATGATGTCGCT	3018		CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).		
Db	861	IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg	880		CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.		
Qy	3019	CTTTCTGCGAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGCAT	3078		CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.		
Db	881	ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp	900		CC EMBL; A5002293; AAF39132.1; -; Genomic DNA.		
Qy	3079	CAAGCTCGGCGATTCAGAGCGGACATCTCTCAATCTATCAGTCTCTGTTGGAGTCAAG	3138		DR PIR; H81722; H81722.		
Db	901	GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys	920		DR InterPro; IPR005546; Auto_transptbeta.		
Qy	3139	TTTGATCGATGTTCTAGTACATCATCTTAATAATATAGCTTTATGGCGCTTATCTGCT	3198		DR InterPro; IPR006315; Autotransporter.		
Db	921	PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys	940		DR InterPro; IPR011427; ChlamPMP_M.		
Qy	3199	GATGTTATCCACCATCTCTGGTACTGAGCAACGCTCTCTATCCCATCAAGAGACATGG	3258		DR InterPro; IPR003368; Chlamydia_PMP.		
Db	941	AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp	960		DR Pfam; PF03797; Autotransporter_1.		
Qy	3259	ACAACAGATGCTTTTATTAGCAAGACATGAGTTGGTTAGAGATCTATGATGCT	3318		DR Pfam; PF02415; Chlam_PMP; 9.		
Db	961	ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla	980		DR Pfam; PF07548; ChlamPMP_M; 1.		
Qy	3319	TCTCTAACAGTATATAGAGTATATGCGCATGGAAGATATGATATCGAGATCTCTCT	3378		DR TIGRFAMS; TIGR01414; autotrans_bar1; 1.		
					DR TIGRFAMS; TIGR01376; POMP repeat; 5.		
					KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.		
					FT SIGNAL 1 25 Potential.		
					FT CHAIN 26 987 Probable outer membrane protein pmpG.		
					SQ SEQUENCE 987 AA; 104867 MW; 32079BD6BBE2DA42 CRC64;		
				Alignment Scores:			
				Pred. No.:	1.28e-246	Length:	987
				Score:	3846.00	Matches:	739
				Percent Similarity:	82.9%	Conservative:	106
				Best Local Similarity:	72.5%	Mismatches:	136
				Query Match:	48.8%	Indels:	38
				DB:	1	Gaps:	13
				US-10-701-844-1 (1-4435) x PMPG_CHLMU (1-987)			
Qy	379	GTGATGCAACGCTTTTCCATAAGTCTTTCTTCAATGATCTAGCTTATCTTCTGCTGC	438				
Db	1	MetMetGlnThrProPheHisLysPhePheLeuAlaMetLeuSerTyr	17				

709 AlaTyrCysArgGlyLeuTrpIleSerGlyIleSerAsnPhePheTyrHisAspGlnAsp 728
 2641 GCTTTAGGTCAGGATATCGGTATATATAGTGGGGGTTATTCCTTAGGACCAATCCTTAC 2700
 729 AlaLeuGlyGlnGlyTyrArgHisIleSerGlyGlyTyrSerIleGlyAlaAsnSerTyr 748
 2701 TTTGGATCATCGATGTTGGTCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTAT 2760
 749 PheGlySerSerMetPheGlyLeuAlaPheThrGluThrPheGlyArgSerIleAspTyr 768
 2761 GTATGTGTGCTGCTTCCATCATCATGCTTCATAGATCCGTTTATCTATCTACCCACAA 2820
 769 ValValCysArgSerAsnAspHisThrCysValGlySerValTyrLeuSerThrArgGln 788
 2821 GCTTTATGTGGATCTTATTTGTCGGAGATGCGTTTATCCGTGCTAGCTACGGGTTGG 2880
 789 AlaLeuCysGlySerCysLeuPheGlyAspAlaPheValArgAlaSerTyrGlyPheGly 808
 2881 AATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGATGCTGCTGGGATAAT 2940
 809 AsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAsnValArgTyrAspAsn 828
 2941 AACTGTCCTGCTCGAGATGTCGAGCGGATACCGATTGCTGATTACTCCATCTAAGCTC 3000
 829 AsnCysValValGlyGluValGlyAlaGlyLeuProIleMetLeuAlaAsnSerLysLeu 848
 3001 TATTTGAATGAGTTGGTCTCTTCGTCGCAAGCTGAGTTTCTTATCCGATCATGAATCT 3060
 849 TyrLeuAsnGluLeuArgProPheValGlnAlaGluPheAlaTyrAlaGluHisGluSer 868
 3061 TTTCAGAGAGAGCGATCAGCTCGGCGATTCAAGCGGACATCTCTTAATCTATCA 3120
 869 PheThrGluArgGlyAspGlnAlaArgGluPheLysSerGlyHisLeuMetAsnLeuSer 888
 3121 GTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTT 3180
 889 IleProValGlyValLysPheAspArgCysSerSerLysHisProAsnLysTyrSerPhe 908
 3181 ATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACAAAGCTCCTA 3240
 909 MetGlyAlaTyrIleCysAspAlaTyrArgSerIleSerGlyThrGluThrThrLeuLeu 928
 3241 TCCCATCAAGACATGACGACAGATGCTTCTTATTTACCAAGACATGAGTGTGGTT 3300
 929 SerHisLysGluThrTrpThrAspAlaPheHisLeuAlaArgHisGlyValMetVal 948
 3301 AGAGGATCTATGCTGCTCTTCAACAGTAATATAGATATATGATGATGATGATGATGAT 3360
 949 ArgGlySerMetTyrAlaSerLeuThrGlyAsnIleGluValTyrGlyHisGlyLysTyr 968
 3361 GAGTATCGAGATGCTTCTCGAGGCTATGTTTGTAGTGCAGAGATGATGATGATGATGAT 3417
 969 GluTyrArgAspAlaSerArgGlyTyrGlyLeuSerIleGlySerLysIleArgPhe 987

RESULT 3
 Q823X5_CHLCV
 ID Q823X5_CHLCV PRELIMINARY; PRT; 1011 AA.
 AC Q823X5_2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymorphic outer membrane protein G family
 DE protein/autotransporter.
 GN OrderedLocusNames=CAA00278;
 OS Chlamydia caviae.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OC NCBI_TaxID=83557;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GPIC;
 RX MEDLINE=22563155; PubMed=12682364; DOI=10.1093/nar/gkg321;
 RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
 Heidelberg J.F.; Holtzapfel E.K., Khouri H.M., Federova N.B.,

RA Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
 RA White O., Salzberg S.L., Haia R.-C., McClarty G., Rank R.G.,
 RA Bavoil P.M., Fraser C.M.;
 RT "Genome sequence of Chlamydia psittaci GPIC":
 RT examining the role of niche-specific genes in the evolution of the
 RT Chlamydiaceae.";
 RL Nucleic Acids Res. 31:2134-2147 (2003).
 DR EMBL; AB016995; AAP05029.1; -; Genomic_DNA.
 DR TIGR; CCA00278; -;
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Auto transportbeta.
 DR InterPro; IPR011427; ChlamPMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07548; ChlamPMP_M; 1.
 DR Pfam; PF02415; Chlam_PMP; 7.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
 KW Complete proteome.
 SQ SEQUENCE 1011 AA; 107952 MW; EBD9A9F1C46E8E6 CRC64;

Alignment Scores:
 Pred. No.: 9,73e-115 Length: 1011
 Score: 1860.50 Matches: 426
 Percent Similarity: 55.2% Conservative: 168
 Best Local Similarity: 39.6% Mismatches: 353
 Query Match: 129
 DB: 27 Gaps:

US-10-701-844-1 (1-4435) x Q823X5_CHLCV (1-1011)
 QY 382 ATGCAACGCTCTTCCATTAAGTCTTCTTCTTCA-----ATGATCTAGCTTATCTTGC 435
 DB 1 MetLysAlaSerLeuArgLysPheLeuIleSerThrThrLeuThrLeuProTyrSerPhe 20
 QY 436 TGTCTCTTAAATGGGGGGATATGCGACGAAATCATGTTCTCTCAAGGAATTTACGAT 495
 DB 21 -----GlnAlaPheSerLeuGluValValProAsnGlyThrThrAsp 35
 QY 496 GGGGAGACGTTAACTGTATCTTCCCTATCTTATAGGAGATCGAGTGGGACTACT 555
 DB 36 Gly--AsnLeuArgGluThrPheProTyrThrIleThrSerAsnProGluGlyThrThr 54
 QY 556 GTTTTTCTGCGAGGAGTTAACTTTAAAAATCTTGACAAATCTTATGTCAGCTTTCCT 615
 DB 55 AlaIleLeuSerGlyAsnLeuAsnLeuLeuAsnLeuAsnSerMetValAlaThrPro 74
 QY 616 TTAAGTGTGTTGGGAACTTATTAGGAGTGTCTTCTAGGAGAGGAGGAGGAGTCTGTTG 675
 DB 75 SerSerCysPhePheAsnSerAlaGlySerMetThrIleValGlyArgAsnHisAsnLeu 94
 QY 676 ACTTTCGAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAAT----- 723
 DB 95 ThrPheThrAsnLeuArgThrSerAlaAsnGlyAlaAlaLeuSerSerIleProThrThr 114
 QY 724 AGCGTCTGATGAGCTGTTTACTTATGAGGGTGTAAAGAAATATCTTTTCAATTCG 783
 DB 115 ThrProGluSerPheProTyrThrIleLysGlyValAsnThrLeuSerPheSerAsnCys 134
 QY 784 AATTCATTACTTGGCTGCTGCTGCTGCAACGACCTAATAAGGTGAGGAGCTCCGACG 843
 DB 135 -----LeuAlaLeuMetAlaArgThrThrAlaProAsnThrThrThrProVal 151
 QY 844 ACAACATCTACACGCTCTAATGCTTACTTATTCTTAAACACAGATCTTTTGTTCATCA 903
 DB 152 -----AsnProAsnGlyGlyAlaPheTyrSerLysAlaProValPheLeuGluAsn 168
 QY 904 AATGAGAGTCTCTATTCTATAGTAATTTAGTCTCTGAGATGGGAGGAGCTATAGTCT 963
 DB 169 IleGlnAsnValLeuPheLysAsnAsnArgAlaAlaAspSerGlyGlyGlyLeuTyrVal 188
 QY 964 AAGAGCTTAAGGTTCAAGGAATTAGCAAGCTTGTCTCTCCAGAAATACTGCTCAA 1023


```

QY 3010 GAGTTGGCTCTTCTGCTCAAGCTGAGTTTCTTATGCGGATCAATCTTTTACAGAG 3069
DB 876 GlnIleIleProPheMetAsnValGlnLeuGlyTyraIaGluHisGlySerPheLysGlu 895
QY 3070 GAAGGCGATCAAGCTCGGCGATCTCAAGACGGGACATCTCTAAATCTATCACTTCTGTT 3129
DB 896 LysLeuAlaGluAlaArgSerPheCysSerArgLeuLeuAsnLeuAlaValProVal 915
QY 3130 GAGTGAAGTTGATCGATGTTCTAGTACATCACTCTTAATAATATAGCTTTTGGCGCT 3189
DB 916 GlyPheLysIleAspArgSerHisSerHisProAspPheTySerLeuAlaIleSer 935
QY 3190 TATATCTGATGCTTATCGCACCATCTCTGCTACTGACACGCTCTCTATCCCATCA 3249
DB 936 TyriIleProAspValTrpArgAsnProGlyCysAsnThrLeuLeuAlaAsnGly 955
QY 3250 GAGACATGACACACATGCTTTCATTTAGCAGACATGAGATGTTGGTTAGAGATCT 3309
DB 956 ValArgTrpLysThrProAlaThrAsnLeuAsnArgHisGlyLeuLeuMetGlnGlySer 975
QY 3310 ATGTATGCTCTCTAACAGTAAATATAGAAATATAGGCGCATGGAAGATATGATATCGA 3369
DB 976 ThrHisThrAlaValLeuSerAsnIleGluIlePheSerHisGlySerCysGluLeuArg 995
QY 3370 GATGCTCTCGAGGCTATGTTTGTAGTGCAGGAGTAGAGTCCGGTTC 3417
DB 996 SerSerArgAsnTyraAsnIleAsnValGlySerLysIleArgPhe 1011

```

RESULT 4

Q5L6J7_CHLAB

```

ID Q5L6J7_CHLAB PRELIMINARY; PRT; 1024 AA.
AC Q5L6J7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polymorphic outer membrane protein.
GN Names:pmprG; Ordered locus names: CAB269;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S2673;
RX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Livingston M., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640(2005).
DR EMBL; CR848038; CAH63725.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 1024 AA, 108663 MW, 26861415C8PD434 CRC64;

```

Alignment Scores:

```

Pred. No.: 3.43e-109 Length: 1024
Score: 1777.00 Matches: 415
Percent Similarity: 53.5% Conservative: 169
Best Local Similarity: 38.0% Mismatches: 362
Query Match: 22.5% Indels: 146
DB: 2 Gaps: 28

```

US-10-701-844-1 (1-4435) x Q5L6J7_CHLAB (1-1024)

```

QY 382 ATGCAACAGCTTTTCCATAGTCTTCTTTCATAGTCTAGCTATTCTTGTCTCT 441
DB 1 MetLysAlaSerPheArgLysPheLeuValSerThrThrLeuThrLeuProCys----- 18
QY 442 TTAATGGGGGGGATATGACAGCAAGAAATCATGTCTCAAGGAATTTACAGGGGAG 501
DB 19 ---SerPheGlnAlaPheSerLeuGluLeuValProAsnGlyThrTyraAsnGlyAsp 37

```

```

QY 502 ACGTTAACTGTATCATTTCCCTATCTGTATAGGAGATCCGAGTGGAGCTACTCTTTT 561
DB 38 ---LeuArgGluMetPheProTyThrIleThrSerAsnProGluGlyThrThrAlaIle 56
QY 562 TCTGAGGAGAGTTAACTTAAATAATCTTGCAATCTTATTCGACGCTTTCCTTAAGT 621
DB 57 LeuSerGlyAsnLeuAsnIleLeuAsnLeuAspAsnSerMetAlaAlaThrProSerSer 76
QY 622 TGTTTGGGAACTTATTAGGAGGTTTACTGTGTTTAGGAGAGGACACTCGTTGTGACTTTC 681
DB 77 CysPhePheAsnSerAlaGlySerMetThrIleIleGlyLysGlyHisAspValThrPhe 96
QY 682 GAGAACATACGACACTTCTACAAATCGGCGAGCTCTAAGT-----AATAGCGCT 729
DB 97 ThrAsnLeuArgThrSerValAsnGlyAlaAlaLeuSerSerLeuIleThrSerSerPro 116
QY 730 GCTGATGAGCTGTTTACTATTGAGGGTTTTAAAGAATTTATCTTTTCCATTCATTCAT 789
DB 117 GluSerPheSerTyThrIleThrGlyValLysThrPheSerCysSerAsnCysSerAla 136
QY 790 TTACTTGGCGTACTGCTCTCAACGACTAATAAGGGTAGCCAGACTCCGACGACAA 849
DB 137 LeuLeu-----GlyArgAsnAsnSerAsnSerIle 146
QY 850 TCTACACCGCTCTAATGGTACTATTATTCTTAAACAGATCTTTTGTACTCAATAATGAG 909
DB 147 LeuSerProLysGlyGlyAlaValTySerLysSerProIlePhePheLysAsnIleGln 166
QY 910 AAGTCTCTCATTTATAGTAAATTTAGTCTCTGAGATGCGGGAGCTATAGATGCTAAGAC 969
DB 167 AsnLeuIlePheLysAspAsnCysAlaAlaAspAsnGlyGlyAlaLeuTrpGlyGlnVal 186
QY 970 TTAACGGTTCAAGGAATATAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGAT 1029
DB 187 ValAspIleSerAsnValThrLysSerLeuLysPheLeuSerAsnValGly---AlaAsn 205
QY 1030 GGGGAGCTTGTCAAGTAGTCTACAGTCTTCTCTGCTATGGCTAAACGAGGCTCTCT---ATT 1086
DB 206 GlyGlyAlaIleGlyAlaSerThrSerLeuAsn-----ValThrArgCysProSerIle 223
QY 1087 GCCTTTGTAGCGAATGTTTCAGGAGTAGAGGGGAGGAGGATGCTGCTGTTCCAGGATGGG 1146
DB 224 LeuPheArgSerAsnSerAlaSerArgLeuGlyGlyAlaIle----- 237
QY 1147 CAGCAGGAGGTGTCATCTACTTCTCAACAGAGATCCA----- 1185
DB 238 ---HisSerValAsnProGlnThrProProProProProGlyAsnGlyValIle 256
QY 1186 -----GTAGTAAGTTTTCAGAAATCTGCGGTAGAGTTTGTATGGGAACGTAGCCCGA 1239
DB 257 AsnThrValValAsnPheSerAspAsnGlySerValGlnPheAspSerAsnAsnAlaLys 276
QY 1240 GTAGGAGGAGGATTTACTCTCTACCGGAACGTTGCTTTCTGTAATATGAAACCTTG 1299
DB 277 SerGlyGlyAlaIleTySerLysGlyAsnIleAsnPheSerAsnValGlnLeuLeu 296
QY 1300 TTTCTCAACATGTTGCTTCTCTCTCTGTTTACATTGCTGCTAAGCAACCAACAGTGGACAG 1359
DB 297 MetGlnAsnAsnThrAlaSerPro-----Glu 305
QY 1360 GCTTCTTAATACGATAATAATTACGGAGTAGGAGGAGCTATCTTCTGTAAGAATGGTGGC 1419
DB 306 IleGlyAspSerAsnGluValLeuGlyGlnGlyAlaIleTyCys-----Thr 322
QY 1420 CAAGCAGGATCCCAATAACTCTGATCAGTTTCTCTTGTATGGAGGGAGGTA----- 1470
DB 323 GlnAlaThrAlaThrLysLysAlaAlaProValPheThrGlyLeuThrIleThrAsnGln 342
QY 1471 -----GTTTTCTTTAGTACATGATGCTGCGGGAACGGGGAGCTATTTATGCCAAA 1524
DB 343 HisAspIlePhePheAlaAsnAsnPheAlaAlaAsnAlaGlyGlyAlaIleTyGlyGlu 362

```


DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Probable outer membrane protein pmpH precursor (Polymorphic membrane
 DE protein H).
 GN Name=pmpH; OrderedLocusNames=CT872;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=D/UW-3/Cx;
 RX MEDLINE=9900809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis.";
 RL Science 282:754-759(1998).
 CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AR001360; AAC68470.1; -; Genomic_DNA.
 DR PIR; H71460; H71460.
 DR PHCI-2DPAGE; O84880; -.
 DR InterPro; IPR005546; Auto_transptbeta.
 DR InterPro; IPR011427; ChlamPMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02415; Chlam_PMP; 6.
 DR Pfam; PF07548; ChlamPMP_M; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
 DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
 KW SIGNAL 1 24 Potential.
 FT CHAIN 25 1016 Probable outer membrane protein pmpH.
 FT SEQUENCE 1016 AA; 107905 MW; E691912C3A2BD6F7 CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 1,05e-96 Length: 1016
 Score: 1589.00 Matches: 311
 Percent Similarity: 92.7% Conservative: 5
 Best Local Similarity: 91.2% Mismatches: 10
 Query Match: 20.2% Indels: 15
 DB: 1 Gaps: 2
 US-10-701-844-1 (1-4435) x PMPH_CHLTR (1-1016)
 QY 3451 ATGCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTGTTCTTATTCG 3510
 DB 1 MetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer 20
 QY 3511 TATGGATTCGCGAGCTCTCTCAAGTGTAAACGCCTAATCAACCACTCTTTTAAGGGA 3570
 DB 21 TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrProPheLysGly 40
 QY 3571 GAGCATGTTTACTTGAATGAGACTGCGCTTTTGTCTCATGCTCTATGCGAGGCTGAGAA 3630
 DB 41 AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn 60
 QY 3631 GGTTCGATTATCTCAGCTAATGCGCAATTTAACGATTACCGACAAACCACTACATTA 3690
 DB 61 GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu 80
 QY 3691 TCATTTACAGATTCCTAAGGCCAGTTCTCTCAAAATTTATGCTTCAATTCAGCAGGAGAG 3750
 DB 81 SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu 100
 QY 3751 ACATTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAAATGTTTCTTCGCGGAGAA 3810

DB 101 ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu 120
 QY 3811 AAGGGATGATCTCCGGGAAAACCGTAGTATTTCCGAGCAGGCGAAGTGCATTTCTGG 3870
 DB 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrp 140
 QY 3871 GATAACTCCGTCGGGTATTTCTCTTATCTACTGTGCCAACCTCATCATCAATCCGGCT 3930
 DB 141 AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro 160
 QY 3931 GCT-----CCACAGTTAGTATGCTCGG 3954
 DB 161 AlaProAlaProAlaProAlaSerSerSerLeuSerProThrValSerAspAlaArg 180
 QY 3955 AAAAGGCTATTTTTCGTGAGACTAGTTTGGAGATCTCAGGCGTCAAAAAAGGGTC 4014
 DB 181 LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysGlyVal 200
 QY 4015 ATGTCGATAATAATCCGCGAATTTCCGAACAGTTTTTCGAGGTPAAGAATAATAAT 4074
 DB 201 MetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsn 220
 QY 4075 GCTGCTGTCGAGGAGCGAGTTCGCTTACCATCAAGTACGACTTTTACAGTTAAAAA 4134
 DB 221 AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs 238
 QY 4135 CTGTAAGGGAAGTTTCTTTCACAGATAAGTACCTCTTCGCGAGCGGAGTGTGTTA 4194
 DB 238 nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyValValTy 258
 QY 4195 TAAAGGCATTGTGCTTTTCAAGACAATGAAGGAGGCATATTTCTCCGAGGGAACACAG 4254
 DB 258 rLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAl 278
 QY 4255 ATACGATGATTTAAGGATTTCTGCTACTAATCAGGATCAGATACGAGACGAGGAG 4314
 DB 278 atyAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyG 298
 QY 4315 CGCTGGAGGAGTATTTGCTCTCCAGATGATCTCTGTAAGTCTTGAAGGCAATAAAGTTC 4374
 DB 298 yGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySe 318
 QY 4375 TATGTTTTTGTATTCAACTTTGCAAAAGCAGAGCGGAAGCATCTTAACGAAAGATT 4434
 DB 318 rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh 338
 QY 4435 C 4435
 DB 338 e 338
 RESULT 6
 Q84FUT7 CHLTR PRELIMINARY; PRT; 981 AA.
 AC Q84FUT7;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymorphic membrane protein H (fragment).
 GN Name=pmpH;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22483673; PubMed=12595433;
 RX DOI=10.1128/IAI.71.3.1200-1208.2003;
 RX Stothard D.R., Toth G.A., Basteiger B.E.;
 RA "Polymorphic membrane protein H has evolved in parallel with the three
 RA disease-causing groups of Chlamydia trachomatis.";
 RL Infect. Immun. 71:1200-1208(2003).
 DR EMBL; AY184168; AAC030032.1; -; Genomic_DNA.
 DR GO; GO:0019867; C:outer membrane; IEA.

DR InterPro; IPR005546; Auto transpetheta.
 DR InterPro; IPR011427; Chlam_PMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR InterPro; IPR002016; Peroxidase.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07548; Autotransporter; 1.
 DR Pfam; PF02415; Chlam_PMP; 1.
 DR Pfam; PF02415; Chlam_PMP; 6.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 981 AA; 104559 MW; 83C957685E57EC10 CRC64;

Alignment Scores:
 Pred. No.: 3.51e-95 Length: 981
 Score: 1566.00 Matches: 303
 Percent Similarity: 99.7% Conservative: 0
 Best Local Similarity: 99.7% Mismatches: 0
 Query Match: 19.9% Indels: 1
 DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x Q84F07_CHLTR (1-981)

QY 3526 TCTCTCAAGTGTAAACGCTTAATGTACCACTCTCTTTAAGGAGAGAGATGTTTACTTG 3585
 DB 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyrlu 20

QY 3586 AATGGAGACTCGCTTTTGTCAATGTCTATCGAGAGCTGAAGAAGGTTCGATTATCTCA 3645
 DB 21 AsnGlyAspCysAlaPheValAsnValTyrlaGlyAlaGluGlySerIleIleSer 40

QY 3646 GCTAATGGGCAATTAACGATTACCGGACAAACCATATCATTTACAGATTCT 3705
 DB 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60

QY 3706 CAAGGGCCAGTCTTCAAAATATGCTTCATTCAGAGAGAGAGACTTACTCTCGA 3765
 DB 61 GlnGlyProValLeuGlnAsnTyrlaPheIleIleSerAlaGlyGluThrLeuThrLeuArg 80

QY 3766 GATTTTTCGAGTCTGATCTTCTCGAAAATATGTTTTCGCGAGAGAAAGGGAATGATCTCC 3825
 DB 81 AspPheSerSerLeuMetPheSerIleAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAACCGTAGTATTTCCGGAGCGGAGAGTATTTTCGGAGTAATCTCGTGGGG 3885
 DB 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerValGly 120

QY 3886 TATTCTCTCTTATCTACTGTCCCAACCTCATCATCTCGCTGCTCCCAACAGTTAGT 3945
 DB 121 TyrSerProLeuSerThrValProThrSerSerSerThrProAlaProThrValSer 140

QY 3946 GATGCTCGGAAGGTCATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005
 DB 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160

QY 4006 AAAGGGTCATGTCGATATAATGTCGGGAATTTCCGAACAGTTTTCGAGGTAAGAT 4065
 DB 161 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180

QY 4066 AATATAATGCTGTGTGGAGGAGTGCGGTTCGCTACACCATCAAGTACGCTTTTAC 4125
 DB 181 AsnAsnAsnAlaGlyGlyGlyGlySer-GlySerAlaThrProSerSerThrThrPheTh 200

QY 4126 AGTTAAAACTGTAAAGGAAAGTTTCTTTTCAAGATAACCTAGCTCTTCGGAGGCGG 4185
 DB 200 rValIleAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGly 220

QY 4186 AGTGTTTATAAAGGCATGTGCTTTTCAAGACATGAAGAGGAGCATATTTCTCCGAGG 4245
 DB 220 yValValTyrlsGlyIleValLeuPheLysAspAsnGluGlyIlePhePheArgGly 240

QY 4246 GAACACAGCATACCATGATTTAAGGATTTCTGCTGCTACTTAATCAGGATCAGATACGGA 4305

DB 240 yAsnThrAlaTyrlAspAspLeuArgIleLeuAlaAlaThrAsnGlnAspGlnAsnThrGl 260
 QY 4306 GACAGGAGCGGTGGAGAGTATTGCTCTCCAGATGATTCGTAAAGTTTGAAGGCAA 4365
 DB 260 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs 280
 QY 4366 TAAAGGTTCTATTGTTTGTGATTACAACTTTGCAAAAGGCGAGGCGAGCATCTTAAC 4425
 DB 280 nLysGlySerIleValPheAspTyrlAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 300

QY 4426 GAAAGAATTC 4435
 DB 300 rLysGluPhe 303

RESULT 7
 Q84F06_CHLTR
 ID Q84F06_CHLTR PRELIMINARY; PRT; 980 AA.
 AC Q84F06;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymorphic membrane protein H (Fragment).
 GN Name=bmpH;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22483673; PubMed=12595433;
 RX DOI=10.1128/IAI.71.3.1200-1208.2003;
 RA Stothard D.R., Toth G.A., Basteiger B.E.;
 RT "Polymorphic membrane protein H has evolved in parallel with the three
 disease-causing groups of Chlamydia trachomatis.";
 RL Infect. Immun. 71:1200-1208(2003).
 DR EMBL; AY184169; AAC03033.1; -; Genomic_DNA.
 DR GO; GO:001967; C:outer membrane; IEA.
 DR InterPro; IPR005546; Auto.transpetheta.
 DR InterPro; IPR011427; Chlam_PMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR InterPro; IPR002016; Peroxidase.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07548; Chlam_PMP; 1.
 DR Pfam; PF02415; Chlam_PMP; 6.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 FT NON TER 1
 FT NON TER 980
 SQ SEQUENCE 980 AA; 104432 MW; BC9B7E33BBD04233 CRC64;

Alignment Scores:
 Pred. No.: 3.22e-94 Length: 980
 Score: 1551.50 Matches: 302
 Percent Similarity: 99.3% Conservative: 0
 Best Local Similarity: 99.3% Mismatches: 0
 Query Match: 19.7% Indels: 2
 DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x Q84F06_CHLTR (1-980)

QY 3526 TCTCTCAAGTGTAAACGCTTAATGTACCACTCTCTTTAAGGAGAGAGATGTTTACTTG 3585
 DB 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyrlu 20

QY 3586 AATGGAGACTCGCTTTTGTCAATGTCTATCGAGAGCTGAAGAAGGTTCGATTATCTCA 3645
 DB 21 AsnGlyAspCysAlaPheValAsnValTyrlaGlyAlaGluGlySerIleIleSer 40

QY 3646 GCTAATGGGCAATTAACGATTACCGGACAAACCATATCATTTACAGATTCT 3705
 DB 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60

QY 3706 CAAGGGCCAGTCTTCAAAATATGCTTCATTCAGAGAGAGAGACTTACTCTCGA 3765
 DB 61 GlnGlyProValLeuGlnAsnTyrlaPheIleIleSerAlaGlyGluThrLeuThrLeuArg 80

QY 3766 GATTTTTCGAGTCTGATCTTCTCGAAAATATGTTTTCGCGAGAGAAAGGGAATGATCTCC 3825
 DB 81 AspPheSerSerLeuMetPheSerIleAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAACCGTAGTATTTCCGGAGCGGAGAGTATTTTCGGAGTAATCTCGTGGGG 3885
 DB 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerValGly 120

QY 3886 TATTCTCTCTTATCTACTGTCCCAACCTCATCATCTCGCTGCTCCCAACAGTTAGT 3945
 DB 121 TyrSerProLeuSerThrValProThrSerSerSerThrProAlaProThrValSer 140

QY 3946 GATGCTCGGAAGGTCATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005
 DB 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160

QY 4006 AAAGGGTCATGTCGATATAATGTCGGGAATTTCCGAACAGTTTTCGAGGTAAGAT 4065
 DB 161 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180

QY 4066 AATATAATGCTGTGTGGAGGAGTGCGGTTCGCTACACCATCAAGTACGCTTTTAC 4125
 DB 181 AsnAsnAsnAlaGlyGlyGlyGlySer-GlySerAlaThrProSerSerThrThrPheTh 200

QY 4126 AGTTAAAACTGTAAAGGAAAGTTTCTTTTCAAGATAACCTAGCTCTTCGGAGGCGG 4185
 DB 200 rValIleAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGly 220

QY 4186 AGTGTTTATAAAGGCATGTGCTTTTCAAGACATGAAGAGGAGCATATTTCTCCGAGG 4245
 DB 220 yValValTyrlsGlyIleValLeuPheLysAspAsnGluGlyIlePhePheArgGly 240

QY 4246 GAACACAGCATACCATGATTTAAGGATTTCTGCTGCTACTTAATCAGGATCAGATACGGA 4305

Db 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuArg 80
QY 3766 GATTTTCGAGTCTGATGTTCTCGAAAATGTTTCTGGGAGAAAAGGGAATGATCTCC 3825
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluGlyMetIleSer 100
QY 3826 GGGAAAACCGTCAGTATTCCCGAGCAGCGAAGTATTCTCGGATAAATCTCCGTGGG 3885
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrPheAsnSerValGly 120
QY 3886 TATTTCTCCTTATCTACTGTGCGAACCTCATCATCACTCCGCTGCTCCAAAGTTAGT 3945
Db 121 TyrSerProLeuSerThrValProThrSerSerSerThrProProAlaProThrValSer 140
QY 3946 GATGCTCGGAAGGTCATATTTTCTGTAGAGCTAGTATTGGAGATCTCAGCGCTCAA 4005
Db 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160
QY 4006 AAAGGGTCATGCTGCATATAATGCGGGGAATTTCCGGAACAGTATTTTCAGGTAAAGAT 4065
Db 161 LysGlyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180
QY 4066 AATAATAATGCTGGTGGAGCGAGTGGGTTCCGCTACACCATCAAGTACGACTTTTAC 4125
Db 181 AsnAsnAlaGlyGlyGlySerGlySerGlySerAlaThrProSerSerThrThrPheTh 200
QY 4126 AGTTAAATCTGTAAAGGAAGTATTTCTTCAAGACATGAGGAGGATATTTCTCCGAGG 4185
Db 200 rValLysAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGly 220
QY 4186 AGTGGTTTAAAGGCAATGCTGTTTCAAGACATGAGGAGGATATTTCTCCGAGG 4245
Db 220 yValValTyrLysGlyIleValLeuPheLysAspAsnGluGlyGlyIlePhePheArg 240
QY 4246 GAACACAGCATACGATGATTAAGGATTTCTGCTGCTACTAATCAGGATCAGAAATAC 4305
Db 240 yAsnThrAlaTyrAspAspLeuArgIleLeu---AlaThrAsnGlnAspGlnAsnThrG 259
QY 4306 GACAGAGCGGPGAGGAGTATTTGCTTCAGATGATTTCTGTAAGTTTGAAGCGAA 4365
Db 259 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAs 279
QY 4366 TAAAGTCTTATGTTTGTATTAACACTTGCATAAGGAGGAGGAGGAGGAGGAGGAGG 4425
Db 279 nLysGlySerIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 299
QY 4426 GAAAGAAATC 4435
Db 299 rLysGluPhe 302

RESULT 8

Q84F08 CHLTR
ID Q84F08 PRELIMINARY; PRT; 980 AA.
AC Q84F08
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic membrane protein H (Fragment).
GN Name=pmph;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RA DOI=10.1128/JAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Bateiger B.E.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RT disease-causing groups of Chlamydia trachomatis";
RL Infect. Immun. 71:1200-1208(2003).
RL EMBL; AY184167; AAC030031.1; -; Genomic_DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto_transpbeta.

DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
FT NON_TER 1 980
SQ SEQUENCE 980 AA; 104460 MW; 27F6C8EA9F43B20P CRC64;

Alignment Scores:

Pred. No.: 3,22e-94 Length: 980
Score: 1551.50 Matches: 302
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 0
Query Match: 19.7% Indels: 2
DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x Q84F08_CHLTR (1-980)

QY 3526 TCTCTCAAGTGTAAACGCTTAATGTAACCACTCTCTTTAAGGAGACGATGTTACTTG 3585
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValTyrLeu 20
QY 3586 AATGAGAGTGGCGTCTTTGTCAATGCTATGAGGAGGCTGAAAGAGGTTTCGATTATCTCA 3645
Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluGluGlySerIleIleSer 40
QY 3646 GCTAATGGCGCAATTTAAAGTAAACCGGACAAAACCATACATTTATCATTTACAGATTCT 3705
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60
QY 3706 CAAGGGCGATGTTCTCAAAATATGCTTCAATTCAGCAGGAGGACACTTACTCTGAGA 3765
Db 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuArg 80
QY 3766 GATTTTCGAGTCTGATGTTCTCGAAAATGTTTCTGGGAGAAAAGGGAATGATCTCC 3825
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100
QY 3826 GGGAAAACCGTCAGTATTCCCGAGCAGCGAAGTATTCTGGGATAACTCCGTGGGG 3885
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrPheAsnSerValGly 120
QY 3886 TATTTCTCCTTATCTACTGTGCGAACCTCATCATCACTCCGCTGCTCCAAAGTTAGT 3945
Db 121 TyrSerProLeuSerThrValProThrSerSerSerThrProProAlaProThrValSer 140
QY 3946 GATGCTCGGAAGGTCATATTTTCTGTAGAGCTAGTATTGGAGATCTCAGCGCTCAA 4005
Db 141 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 160
QY 4006 AAAGGGTCATGTTCCGATAAATGCGGGGAATTTCCGGAACAGTATTTTCAGGTAAAGAT 4065
Db 161 LysGlyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyLysAsn 180
QY 4066 AATAATAATGCTGGTGGAGCGAGTGGGTTCCGCTACACCATCAAGTACGACTTTTAC 4125
Db 181 AsnAsnAlaGlyGlyGlySerGlySerAlaThrProSerSerThrThrPheTh 200
QY 4126 AGTTAAATCTGTAAAGGAAGTATTTCTTCAAGACATGAGGAGGATATTTCTCCGAGG 4185
Db 200 rValLysAsnCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGly 220
QY 4186 AGTGGTTTAAAGGCAATGCTGTTTCAAGACATGAGGAGGATATTTCTCCGAGG 4245
Db 220 yValValTyrLysGlyIleValLeuPheLysAspAsnGluGlyGlyIlePhePheArg 240
QY 4246 GAACACAGCATACGATGATTAAGGATTTCTGCTGCTACTAATCAGGATCAGAAATAC 4305
Db 240 yAsnThrAlaTyrAspAspLeuArgIleLeu---AlaThrAsnGlnAspGlnAsnThrG 259

```

QY 4306 GACAGGAGCGGTGAGAGGATTTTGTCTCCAGATGATTTCTGTAAAGTTTGAAGGCAA 4365
Db 259 uThrGlyGlyGlyGlyValIleCySerProAspSerValIlePheGluGlyAs 279
QY 4366 TAAAGGTTCTATTGTTTGTGATTCAACTTTGCAAAAGCGAGAGCGGAAGCCTTAAC 4425
Db 279 nLyseGlySerIleValPheAspTyAsnPhAlaLysGlyArgGlyGlySerIleLeuTh 299
QY 4426 GAAAGAATTC 4435
Db 299 rLyseGluPhe 302

RESULT 9
Q83U76_CHLTR
ID Q83U76_CHLTR PRELIMINARY; PRT; 989 AA.
AC Q83U76;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Polymorphic membrane protein H (Fragment).
GN Name=pmpH;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RX DOI=10.1128/JAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Basteiger B.E.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RT disease-causing groups of Chlamydia trachomatis.";
RL Infect. Immun. 71:1200-1208(2003).
DR EMBL; AY184165; AAO30029.1; -; Genomic_DNA.
DR EMBL; AY184166; AAO30030.1; -; Genomic_DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; AutoTranspBeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01376; POMP repeat; 4.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 989
SQ SEQUENCE 989 AA; 104934 MW; 75D5EAAF4633F691 CRC64;

Alignment Scores:
Pred. No.: 4.5e-88 Length: 989
Score: 1459.00 Matches: 287
Percent Similarity: 93.0% Conservative: 5
Best Local Similarity: 91.4% Mismatches: 9
Query Match: 18.5% Indels: 13
DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x Q83U76_CHLTR (1-989)
QY 3526 TCTCTCAAGTGTAAACGCTTAAGTACCACTCTTTTAAAGGAGAGGATGTTTACTTG 3585
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspValTyrlLeu 20
QY 3586 AATGAGACTCGCTTTTGTCTAATGTCATATGAGGAGTGAAGAGGTTCCATTATCTCA 3645
Db 21 AsnGlyAspCysAlaPheValAsnValTyrlAlaGlyAlaGluAsnGlySerIleLeuSer 40
QY 3646 GCTAATGCGGCAATTAAACGATTACCGGACAAACCATACATTATCATTTACAGATTCT 3705
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60
QY 3706 CAAGGCCAGTCTCTCAAAATATGCTTCTATTCAGGAGGAGACACTTACTCTGAGA 3765

```

```

Db 61 GlnGlyProValLeuGlnAsnTyrlAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80
QY 3766 GATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCTTCGAGAGAAAGGAATGATCTCC 3825
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100
QY 3826 GCGAAACCTGAGTATTTCGAGAGCGAGGAGAGTGAATTTCTCGGATTAATCTCGTGGGG 3885
Db 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrpAspAsnSerValGly 120
QY 3886 TATTTCTCTTTATCTACTGTGCCAACCTCATCATCACTCCGCTGCT----- 3933
Db 121 TyrSerProLeuSerThrValProAlaSerThrProThrProAlaProAlaProAla 140
QY 3934 -----CCACAGTTAGTGTCTCGGAAAGGCTGCTATTTTCTGTGTA 3975
Db 141 AlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePheSerVal 160
QY 3976 GAGCTAGTTTGGAGATCTCAGCGGTCAAAAAGGGGTGATGTCGATATATATGCCGGG 4035
Db 161 GluThrSerLeuGluIleSerGlyValLysGlyGlyValMetPheAspAsnAlaGly 180
QY 4036 AATTTCCGAACAGTTTTCGAGGTAAAGATAATAAATCTGCTGCTGAGGAGGAGTGGG 4095
Db 181 AsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGlySer-Gl 200
QY 4096 TTCGCTACACCATCAAGTACGACTTTTACAGTTTAAAAAAGTAAAGGAAAGTTTCTTT 4155
Db 200 YSerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLysValSerPh 218
QY 4156 CACAGATAAGTAGCTCTTCGCGAGCGGAGGTGTTTATAAAGGCATGTCGCTTTTCAA 4215
Db 218 eThrAspAsnValAlaSerCysGlyGlyGlyValValTyrlLysGlyThrValLeuPheLy 238
QY 4216 AGACAATGAAGGAGGCATATTTCTTCGAGGGAACACAGCATACGATCATTTAAGGATTCT 4275
Db 238 sAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrlAspAspLeuGlyIleLe 258
QY 4276 TGCTGTCTACTAATCAGGATCAGAAATACGAGACAGGAGCGGTGGAGAGGTTATTTGCTC 4335
Db 258 uAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyGlyValIleCysSe 278
QY 4336 TCCGAGATGATCTGTAAGTTTGAAGCAATAAAGTTCTATTGCTTTTGTGATTAACAATT 4395
Db 278 rProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAspTyrlAsnPh 298
QY 4396 TGCAAAAGCGAGGCGGAAGCATCTTAACGAAAGAAATTC 4435
Db 298 eAlaLysGlyArgGlySerIleLeuThrLysGluPhe 311

RESULT 10
Q84F09_CHLTR
ID Q84F09_CHLTR PRELIMINARY; PRT; 989 AA.
AC Q84F09;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic membrane protein H (Fragment).
GN Name=pmpH;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RX DOI=10.1128/JAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Basteiger B.E.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RT disease-causing groups of Chlamydia trachomatis.";
RL Infect. Immun. 71:1200-1208(2003).
DR EMBL; AY184164; AAO30028.1; -; Genomic_DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; AutoTranspBeta.

```

```
DR InterPro; IPR011427; ChlamPMP_M.  
DR InterPro; IPR003368; Chlamydia_PMP.  
DR Pfam; IPR002016; Peroxidase.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF07548; ChlamPMP_M; 1.  
DR Pfam; PF02415; Chlam_PMP; 6.  
DR TIGRFAMs; TIGR01376; POMP repeat; 4.  
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.  
FT NON_TER 1  
FT NON_TER 989  
SQ SEQUENCE 989 AA; 104953 MW; 75D5F65D4F21E371 CRC64;  
  
Alignment Scores:  
Pred. No.: 4,5e-88 Length: 989  
Score: 1459.00 Matches: 287  
Percent Similarity: 93.0% Conservative: 5  
Best Local Similarity: 91.4% Mismatches: 13  
Query Match: 18.5% Indels: 9  
DB: 2 Gaps: 2  
  
US-10-701-844-1 (1-4435) x Q84FU9_CHLTR (1-989)  
  
QY 3526 TCTCTCAAGTGTAAACGCTAATGTAAACACTCTCTTTAAGGAGACGATCTTACTTG 3585  
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspValTyrLeu 20  
  
QY 3586 AATGGAGACTGCGCTTTCTCAATGTCATGAGGAGCTGAAGAGTTTCGATTATCTCA 3645  
Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleSer 40  
  
QY 3646 GCTAATGGGACAAATTACGATTACCGCAAAACCATACATTATCATTTACAGATTCT 3705  
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60  
  
QY 3706 CAAGGCCAGTCTCTCAAAATATATGCTCTTCATTTACAGAGGAGACACTTACTCTGAGA 3765  
Db 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80  
  
QY 3766 GATTTTTCAGTCTGATGTTCTCGAAAAATGTTTCTTGGGAGAAAAAGGAATGATCTCC 3825  
Db 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100  
  
QY 3826 GGGAAACGTCGATTTTTCGGGAGCAGGAGTGTATTTCTGGGATAACTCCGCGGG 3885  
Db 101 GlyLeuThrValSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerValGly 120  
  
QY 3886 TATTCCTCTTATCTACTGTGCCAACCTCATCATCAACTCCGCTGCT----- 3933  
Db 121 TyrSerProLeuSerThrValProAlaSerThrProThrProAlaProAlaProAla 140  
  
QY 3934 -----CCAACAGTTAGTGTGCTCGGAAAGGGTCTATTTTCTGTA 3975  
Db 141 AlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePheSerVal 160  
  
QY 3976 GAGACTAGTTGAGATCTCAGCGCTCAAAAAAGGGTCAATGTCGATATATATCGCGG 4035  
Db 161 GluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAsnAlaGly 180  
  
QY 4036 AATTCGGACAGTTCCTCGAGGTAAAGTAATAATGCTGCTGTCGTCGAGGAGTCGGG 4095  
Db 181 AsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsnAlaGlySerGlyGlySer-GI 200  
  
QY 4096 TTCGGCTACACCATCAAGTACGACTTTTACAGTTAAAACTGTAAAGGGGAAAGTCTTTT 4155  
Db 200 ySerAlaThrProSer-----PheThrValLysAsnCysLysGlyLysValSerPh 218  
  
QY 4156 CACAGATAACGTAGCTCTTCGGAGCGCGAGTGGTTTATAAGGCATTCGCTTTTCAA 4215  
Db 218 eThrAspAsnValAlaSerCysGlyGlyValValTyrLysGlyThrValLeuPheLys 238  
  
QY 4216 AGCAATGAAGGAGGCATATCTTCGAGGGGMAACACAGCATACGATCATTTAAGGATTCT 4275  
Db 238 BASpAsnGluGlyGlyIlePhePheArgLysAsnThrAlaTyrAspAspLeuGlyIleLe 258
```

```
QY 4276 TGCTGCTACTAATCAGGATCAGAAATACGAGACAGGAGCGGTGAGGAGTATTATGCTC 4335  
Db 258 uAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyGlyValIleCysSe 278  
  
QY 4336 TCCAGATGATTCTGTAAAGTTTGAAGCAATAAAGGTTCTATTGTTTGGATTACAACTT 4395  
Db 278 xProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAspTyrAsnPh 298  
  
QY 4396 TGCAAAAGCGCAGAGCGGAGCATCCTTAACGAAAGAAATTC 4435  
Db 298 eAlaLysGlyArgGlyGlySerIleLeuThrLysGluPhe 311  
  
RESULT 11  
Q83UW2_CHLTR  
ID Q83UW2_CHLTR PRELIMINARY; PRT; 991 AA.  
AC Q83UW2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Polymorphic membrane protein H (Fragment).  
GN Name=pmpH;  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OC NCBI_TaxID=813;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22483673; Pubmed=12595433;  
RX DOI=10.1128/IAI.71.3.1200-1208.2003;  
RA Toth G.A., Batteiger B.E.;  
RT "Polymorphic membrane protein H has evolved in parallel with the three  
disease-causing groups of Chlamydia trachomatis.";  
RL Infect. Immun. 71:1200-1208(2003).  
DR EMBL; AY184159; AAC0023.1; -; Genomic DNA.  
DR EMBL; AY184163; AAC0027.1; -; Genomic DNA.  
DR EMBL; AY184162; AAC0026.1; -; Genomic DNA.  
DR GO; GO:0019867; C:outer membrane; IEA.  
DR InterPro; IPR005546; Auto transpbtbeta.  
DR InterPro; IPR011427; ChlamPMP_M.  
DR InterPro; IPR003368; Chlamydia_PMP.  
DR InterPro; IPR002016; Peroxidase.  
DR Pfam; PF03797; Autotransporter; 1.  
DR Pfam; PF07548; ChlamPMP_M; 1.  
DR Pfam; PF02415; Chlam_PMP; 6.  
DR TIGRFAMs; TIGR01376; POMP repeat; 4.  
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.  
FT NON_TER 1  
FT NON_TER 991  
SQ SEQUENCE 991 AA; 105133 MW; 09586A6137569398 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,31e-87 Length: 991  
Score: 1452.00 Matches: 286  
Percent Similarity: 92.1% Conservative: 5  
Best Local Similarity: 90.5% Mismatches: 10  
Query Match: 18.4% Indels: 15  
DB: 2 Gaps: 2  
  
US-10-701-844-1 (1-4435) x Q83UW2_CHLTR (1-991)  
  
QY 3526 TCTCTCAAGTGTAAACGCTAATGTAAACACTCTCTTTAAGGAGACGATGTTACTTG 3585  
Db 1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspValTyrLeu 20  
  
QY 3586 AATGGAGACTGCGCTTTTCTCAATGTCATGAGGAGCTGAAGAGTTTCGATTATCTCA 3645  
Db 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleSer 40  
  
QY 3646 GCTAATGGGACAAATTACGATTACCGCAAAACCATACATTATCATTTACAGATTCT 3705  
Db 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60  
  
QY 3706 CAAGGCCAGTCTCTCAAAATATATGCTCTTCATTTACAGAGGAGACACTTACTCTGAGA 3765
```

		ProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys	80
Db	61	GInGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys	::::
Qy	3766	GATTTTCGAGTCTCGATTCTCGAAAAATGTTCCTCGCGAGAAAAGGAATCATCTCC	3825
Db	81	AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer	100
Qy	3826	GGGAAACCGTAGATTTCGGAGCAGGCCAAGTAGTTTTCTGGNATAACTCCGTGGGG	3885
Db	101	GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrpAspAsnSerValGly	120
Qy	3886	TATTCTCCTTTATCTACTGTGCCAACCTCATCACTCCGCGTGCT-----	3933
Db	121	TyrSerProLeuSerIleValProAlaSerThrProThrProProAlaProAla	140
Qy	3934	-----CCACAGTTAGTGATCTCGGAAGGCTCTATTTTT	3969
Db	141	ProAlaAlaSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePhe	160
Qy	3970	TCTGTAGAGCTAGTTGGAGTCTCAGCGCTCAAAAAGGGCTCATGTTTCGATATAAT	4029
Db	161	SerValGluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAsn	180
Qy	4030	GCCGGGAATTCGGNAACAGTTTTTCGAGGTAAAGAATAATAATATCTGCTGGTGAGGC	4089
Db	181	AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsnAlaGlySerGlyGly	200
Qy	4090	AGTGGGTTCCGCTACACATCAAGTACGACTTTTACAGTTAAAACTGTAAGGGGAAAGT	4149
Db	201	Ser-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLysVa	218
Qy	4150	TTCTTTCCACAGATAAAGTACCTCTTCGGAGCGGAGTGGTTTATAAAGGCATTGCT	4209
Db	218	IserPheThrAspAsnValAlaSerCysGlyGlyGlyValValTyrLysGlyThrValIle	238
Qy	4210	TTTCAAAGACAATCAAGGAGCATATTCTTCGAGGGGAACACAGCATACATATTAAG	4269
Db	238	uPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLeuGI	258
Qy	4270	GATTCTTCGCTACTTAATCAGGNATCAGATACGGACACAGGCGCGGTGGAGGTTAT	4329
Db	258	yileLeuAlaAlaThrSerArgaspGlnAsnThrGluThrGlyGlyGlyValII	278
Qy	4330	TTGCTCTCCAGATGATTCTGTAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTGATTA	4389
Db	278	eCysSerProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAspTy	298
Qy	4390	CAACTTCGAAAAGCGACAGCGGAAGCATCTTAACGAAAGATTC	4435
Db	298	rAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLvsGluPhe	313

RESULT 12	QY	3934	-----CCAACAGTTAGTGCATCTCGGAAAGGTCATATTTT	3969
Q84FV0 CHLTR	QY	3934	-----CCAACAGTTAGTGCATCTCGGAAAGGTCATATTTT	3969
ID Q84FV0 CHLTR PRELIMINARY; PRT; 991 AA.	Db	141	ProAlaAlaSerSerLeuSerProThrValSerAspAlaArgIysGlySerIlePhe	160
AC Q84FV0;	QY	3970	TCTGTAGAGACTAGTTTCGAGATCTCAGGCGTCAAAAAAGGGTCATGTTTCGATATAAT	4029
DT 01-JUN-2003 (TREMBLrel. 24, Created)	Db	161	SerValGluThrSerLeuGluIleSerGlyValIlysGlyValMetPheAspAsnAsn	180
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)	QY	4030	GCCGGGAATTTTCGGAAACAGTTTTTCGAGGTAAGAATAATAATAATCTCGTGTGAGGC	4089
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)	Db	181	AlaGlyAsnPheGlyThrValPheArgIysSerAsnAsnAlaGlySerGlyGly	200
DE Polymorphic membrane protein H (Fragment).	QY	4090	AGTGGGTTCCGCTACACCATCAAGTACGACTTTTACAGTTTAAAAACTGTAAAGGGAAGT	4149
QN Name=pmpH;	Db	201	Ser-GlySerAlaThrThrProSer-----PheThrValIysAsnCysIysGlyIysVa	218
OS Chlamydia trachomatis.	QY	4150	TTCTTTCACATAACGTAGCTCTTCGGGAGCGGAGTCGTTTATATAAGGCATTGCT	4209
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.	Db	218	IserPheThrAspAsnValAlaSerCysGlyIysGlyValValTyrIysGlyThrValLe	238
NCBI_TaxID=813;	QY	4210	TTTCAAAGCAATGAAGGAGCATATTCTTCCGAGGGGAACACAGCATACCATGATTTAAG	4269
[1]				
NUCLEOTIDE SEQUENCE.				
RP MEDLINE=22483673; PubMed=12595433;				
RX DOI=10.1128/IAI.71.3.1200-1208.2003;				
RA Stothard D.R., Toth G.A., Basteiger B.E.;				
RT Polymorphic membrane protein H has evolved in parallel with the three				
RT disease-causing groups of Chlamydia trachomatis.;				
RL Infect. Immun. 71:1200-1208(2003).				
DR EMBL; AY184161; AAC030025.1; -; Genomic DNA.				
DR GO; GO:0019867; C:outer membrane; IEA_				


```

Db      238 uphelysAspAsnGluGlyGlyLePhePheArgGlyAsnThrAlaTyAspAspLeuGl 258
QY      4270 GATTCTGTCTGCTACTTAATCAGATCAGATACGAGACAGAGCGCGTGGAGGATTAT 4329
Db      258 ylleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyValIi 278
QY      4330 TTGCTCTCCAGATGATTCTGTAAGTTTGAAGCAATAAAGGTTCTATTGTTTGGATTA 4389
Db      278 eCySerProAspAspSerVallysPheGluGlyAsnLysGlySerIleValPheAspTy 298
QY      4390 CAACTTTGCAAAAGGCGAGCGGAGCATCTTAACGAAGAATTC 4435
Db      298 rAsnPheAlaLysGlyArgGlySerIleLeuThrLysGluPhe 313

RESULT 13
Q84FV1_CHLTR
ID      Q84FV1_CHLTR PRELIMINARY; PRT; 991 AA.
AC      Q84FV1;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Polymorphic membrane protein H (fragment).
GN      Name=pmph;
OS      Chlamydia trachomatis.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=813;
[1]
RN      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=22483673; PubMed=12595433;
RA      Stothard D.R., Toth G.A., Basteiger B.E.;
RT      "Polymorphic membrane protein H has evolved in parallel with the three
RL      disease-causing groups of Chlamydia trachomatis.";
DR      EMBL; AY184156; AAO30020.1; -; Genomic DNA.
GO      GO:0019867; C:outer membrane; IEA.
DR      InterPro; IPR005546; Auto transportbeta.
DR      InterPro; IPR011427; ChlamPMP_M.
DR      InterPro; IPR003368; Chlamydia_PMP.
DR      InterPro; IPR002016; Peroxidase.
DR      Pfam; PF03797; Autotransporter; 1.
DR      Pfam; PF02415; ChlamPMP M; 1.
DR      Pfam; PF02415; Chlam_PMP; 6.
DR      TIGRFAMs; TIGR01376; POMP_repeat; 4.
DR      PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
FT      NON_TER 1
FT      NON_TER 991
SQ      SEQUENCE 991 AA; 105126 MW; 0004DC48387A2478 CRC64;

Alignment Scores:
Pred. No.: 1.31e-87 Length: 991
Score: 1452.00 Matches: 286
Percent Similarity: 92.1% Conservative: 5
Best Local Similarity: 90.5% Mismatches: 10
Query Match: 18.4% Indels: 15
DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x Q84FV1_CHLTR (1-991)
QY      3526 TCTCTCAAGTGTTAAAGCCCTAAGTAACCACTCCTTTTAAAGGAGAGCATGTTTACTTG 3585
Db      1 SerProGlnValLeuThrProAsnValThrThrProPheLysGlyAspAspValIleu 20
QY      3586 AATGAGACTCGCTTTTGTCAATGCTCTATCAGAGAGCTGAAGAGGTTTCGATTATCTCA 3645
Db      21 AsnGlyAspCysAlaPheValAsnValTyAlaGlyAlaGluAsnGlySerIleIleSer 40
QY      3646 GCTAATGGCGCAATTTTAAACGATTACCGGACAAACCATATTCATTATTCAGATTCT 3705
Db      41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAspSer 60
QY      3706 CAAGGCCAGTCTTCAAAATATGCTTCATTTTCAGCGAGAGACACTTACTCTGAGA 3765

```

```

61 GlnGlyProValLeuGlnAsnTyAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80
3766 GATTTTTCGAGTCTCATGTTCTCGAAAAATGTTTCTCGGAGAAAAAGGAATGATCTCC 3825
Db      81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100
3826 GGGAAAAACGTCGATGATTTCGCGAGCAGCGCAAGTAGATTTTCTGGGATACTCCGTCGGG 3885
Db      101 GlyLysThrValIleSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerValGly 120
3886 TATTCTCTCTTATCTACTGTGCAACCTCATCATCAACTCCCGCTGCT----- 3933
Db      121 TyrSerProLeuSerIleValProAlaSerThrProThrProProAlaProAlaProAla 140
3934 -----CCAACAGTTAGTAGTCTCGGAAAAAGGGCTATTATTTT 3969
Db      141 ProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIlePhe 160
3970 TCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAAGGGGTCTATGTTCCATATATAT 4029
Db      161 SerValGluThrSerLeuGluIleSerGlyVallysLysGlyValMetPheAspAsnAsn 180
4030 GCGGGAAATTCGGAACAGTTCGAGGTAAAGATAATAATAATCTGCTGCTGTCGAGGC 4089
Db      181 AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsnAlaGlySerGlyGly 200
4090 AGTGGGTTCCGCTACACCATCAAGTAGCAGCTTTTACAGTTAAAAAAGCTGTAAGGGAAAGT 4149
Db      201 Ser-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLysVa 218
4150 TTCTTTTACAGATTAACGTAGCTTTCGCGAGCGCGAGTGGTTTATAAAGGCATTTGTGCT 4209
Db      218 IserPheThrAspAsnValAlaSerCysGlyGlyGlyValValTyLysGlyThrValLe 238
4210 TTTCAAAGCAATGAAGGAGGCGATATCTTCGAGGAGACACAGCATCATGATTTTAAG 4269
Db      238 uPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyAspAspLeuGl 258
4270 GATTCTTGTCTTACTTAATCAGCATCAGATAACGAGACAGAGGCGCGTGGAGGAGTTAT 4329
Db      258 ylleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyGlyValIi 278
4330 TTGCTCTCCAGATGATTCTGTAAGTTTGAAGCAATAAAGGTTCTATTGTTTGGATTA 4389
Db      278 eCySerProAspAspSerVallysPheGluGlyAsnLysGlySerIleValPheAspTy 298
4390 CAACTTTGCAAAAGGCGAGCGGAGCATCTTAACGAAGAATTC 4435
Db      298 rAsnPheAlaLysGlyArgGlySerIleLeuThrLysGluPhe 313

RESULT 14
Q83TJ6_CHLTR
ID      Q83TJ6_CHLTR PRELIMINARY; PRT; 993 AA.
AC      Q83TJ6;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE      Polymorphic membrane protein H (fragment).
GN      Name=pmph;
OS      Chlamydia trachomatis.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=813;
[1]
RN      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=22483673; PubMed=12595433;
RA      Stothard D.R., Toth G.A., Basteiger B.E.;
RT      "Polymorphic membrane protein H has evolved in parallel with the three
RL      disease-causing groups of Chlamydia trachomatis.";
DR      EMBL; AY184156; AAO30020.1; -; Genomic DNA.
DR      EMBL; AY184158; AAO30022.1; -; Genomic DNA.
DR      EMBL; AY184157; AAO30021.1; -; Genomic DNA.

```

DR GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR005546; Auto_transphbeta.
 DR InterPro; IPR011427; Chlam_PMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR InterPro; IPR002016; Peroxidase.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07548; ChlamPMP_M; 1.
 DR Pfam; PF02415; Chlam_PMP; 6.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 FT NON_TER 1
 FT NON_TER 993
 SQ SEQUENCE 993 AA; 105683 MW; D20AABD706DEABA3 CRC64;

Alignment Scores:
 Pred. No.: 1.89e-84 Length: 993
 Score: 1404.50 Matches: 279
 Percent Similarity: 90.9% Conservative: 9
 Best Local Similarity: 88.0% Mismatches: 13
 Query Match: 17.8% Indels: 16
 DB: 2 Gaps: 3

US-10-701-844-1 (1-4435) x Q83TJ6_CHLTR (1-993)

QY 3526 TCCTCTCAAGTGTAAACGCTAATGTAACCACTCTTTTAAAGGAGAGCATGTTACTTGG 3585
 DB 1 SerProGlnValLeuThrProAsnValThrProPheLysGlyAspValTyrLeu 20

QY 3586 AATGAGACTGCGCTTTTGTCAATGCTATGCGAGGCTGAAGAAGGTTCCGATTATCTCA 3645
 DB 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleIleSer 40

QY 3646 GCTAATGCGCAATTAACGATTACCGACAAACCATACATTATCATCATGATTCT 3705
 DB 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrHisSer 60

QY 3706 CAAGGGCCAGTTCTCAAAATATGCTTCATTCAGCAGGAGACACTTACTCTGAGA 3765
 DB 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80

QY 3766 GATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTCGGAGAAAAGGAATGATCTCC 3825
 DB 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAAACGTAAGTATTTCCGGAGACGCGAAGTATTTCTCGGATAACTCCGTGGGG 3885
 DB 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheArgAspAsnSerValGly 120

QY 3886 TATTCCTCTTATCTACTGTGCGAACCTCATCATCACTCCGCTGCT----- 3933
 DB 121 TyrSerProLeuSerThrValProThrSerSerSerThrProAlaProAspProSer 140

QY 3934 -----CCACAGTTAGTGTGCTCGGAAAAGGGTCTATTTT 3969
 DB 141 ProThrAlaSerSerSerSerProThrValSerAspAlaArgGlnGlySerIlePhe 160

QY 3970 TCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAGGGGTGATGTTGATTAATAT 4029
 DB 161 SerIleGluThrSerLeuGluIleSerGlyIleLysLysGluValMetPheAspAsnAsn 180

QY 4030 GCCGGGAATTCGGAACAGTTTTCGAGGTAAGATAATAATGCTGGTGGTGAGGC 4089
 DB 181 AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGly 200

QY 4090 AGTGGGTTCCGCTACACCATCAAGTACGATTTTACAGTTAAAACTGTAAAGGGAAAGT 4149
 DB 201 Ser-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLysVal 218

QY 4150 TTCTTTTCAGATAACGTAAGCTCTTCGGAGGCGAGGTGGTTTATAAAGCATTTGCT 4209
 DB 218 LserPheThrAspAsnValAlaSerCysGlyGlyValValTyrLysGlyThrValIle 238

QY 4210 TTTCAAAGACATGAAGGAGGCATATTCTTCGAGGGAACACAGCATACGATGATTAAAG 4269

DB 238 uPheLysAspAsnGlnGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLeuG1 258
 QY 4270 GATTCTTGCTCTACTAATCAAGATCAG---AATACCGAGACAGAGCGGTGAGGAGT 4326
 DB 258 yIleLeuAlaAlaThrSerGlnAspGlnSerAsnThrGluThrGlyGlyGlyAl 278
 QY 4327 TATTGCTCTCCAGATGATCTGTAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTGA 4386
 DB 278 aileCysSerProGluAspThrValThrPheGluGlyAsnLysGlySerIleValPheAs 298
 QY 4387 TTACAACCTTTGCAAAAGGACAGGCGAAGCATCTTAACGAAGAATTC 4435
 DB 298 pTyrAsnPheAlaLysGlyArgGlyGlyIleLeuThrLysPhe 314

RESULT 15
 Q84FV2_CHLTR
 ID Q84FV2_CHLTR PRELIMINARY; PRT; 993 AA.
 AC Q84FV2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polymorphic membrane protein H (Fragment).
 GN Name=pmpH;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22483673; PubMed=12595433;
 RX DOI=10.1128/JAI.71.3.1200-1208.2003;
 RA Stothard D.R., Toth G.A., Batteiger B.E.;
 RT "Polymorphic membrane protein H has evolved in parallel with the three
 disease-causing groups of Chlamydia trachomatis.";
 RL Infect. Immun. 71:1200-1208(2003).
 DR EMBL; AY184155; AAC0019.1; -, Genomic_DNA.
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR005546; Auto_transphbeta.
 DR InterPro; IPR011427; Chlam_PMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR InterPro; IPR002016; Peroxidase.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07548; ChlamPMP_M; 1.
 DR Pfam; PF02415; Chlam_PMP; 6.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 4.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 FT NON_TER 1
 FT NON_TER 993
 SQ SEQUENCE 993 AA; 105699 MW; 381340DD86400E39 CRC64;

Alignment Scores:
 Pred. No.: 1.89e-84 Length: 993
 Score: 1404.50 Matches: 279
 Percent Similarity: 90.9% Conservative: 9
 Best Local Similarity: 88.0% Mismatches: 13
 Query Match: 17.8% Indels: 16
 DB: 2 Gaps: 3

US-10-701-844-1 (1-4435) x Q84FV2_CHLTR (1-993)

QY 3526 TCCTCTCAAGTGTAAACGCTAATGTAACCACTCTTTTAAAGGAGAGCATGTTACTTGG 3585
 DB 1 SerProGlnValLeuThrProAsnValThrProPheLysGlyAspValTyrLeu 20

QY 3586 AATGAGACTGCGCTTTTGTCAATGCTATGCGAGGCTGAAGAAGGTTCCGATTATCTCA 3645
 DB 21 AsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIleIleSer 40

QY 3646 GCTAATGCGCAATTAACGATTACCGACAAACCATACATTATCATCATGATTCT 3705
 DB 41 AlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrHisSer 60

QY 3706 CAAGGGCCAGTTCTCAAAATATGCTTCATTCAGCAGGAGACACTTACTCTGAGA 3765
 DB 61 GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys 80

QY 3766 GATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTCGGAGAAAAGGAATGATCTCC 3825
 DB 81 AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer 100

QY 3826 GGGAAAACGTAAGTATTTCCGGAGACGCGAAGTATTTCTCGGATAACTCCGTGGGG 3885
 DB 101 GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheArgAspAsnSerValGly 120

QY 3886 TATTCCTCTTATCTACTGTGCGAACCTCATCATCACTCCGCTGCT----- 3933
 DB 121 TyrSerProLeuSerThrValProThrSerSerSerThrProAlaProAspProSer 140

QY 3934 -----CCACAGTTAGTGTGCTCGGAAAAGGGTCTATTTT 3969
 DB 141 ProThrAlaSerSerSerSerProThrValSerAspAlaArgGlnGlySerIlePhe 160

QY 3970 TCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAGGGGTGATGTTGATTAATAT 4029
 DB 161 SerIleGluThrSerLeuGluIleSerGlyIleLysLysGluValMetPheAspAsnAsn 180

QY 4030 GCCGGGAATTCGGAACAGTTTTCGAGGTAAGATAATAATGCTGGTGGTGAGGC 4089
 DB 181 AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGly 200

QY 4090 AGTGGGTTCCGCTACACCATCAAGTACGATTTTACAGTTAAAACTGTAAAGGGAAAGT 4149
 DB 201 Ser-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGlyLysVal 218

QY 4150 TTCTTTTCAGATAACGTAAGCTCTTCGGAGGCGAGGTGGTTTATAAAGCATTTGCT 4209
 DB 218 LserPheThrAspAsnValAlaSerCysGlyGlyValValTyrLysGlyThrValIle 238

QY 4210 TTTCAAAGACATGAAGGAGGCATATTCTTCGAGGGAACACAGCATACGATGATTAAAG 4269

```

Db      61  GlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeuLys  80
QY      3766  GATTTTCGAGTCTGATGTTCTCGAAATGTTTCTCGGAGAGAAAGGATGATCTCC  3825
Db      81  AspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGluLysGlyMetIleSer  100
QY      3826  GCGAAACCGTCAGTATTTCCGAGCAGCGAAGTGATTTCTCGGATAACTCCCTGGGG  3885
Db      101  GlyLysThrValSerIleSerGlyAlaGlyGluValIlePheArgAspAsnSerValGly  120
QY      3886  TATTCTCTCTTACTACTGTCACACCTCATCATCAACTCCGCCTGCT  3933
Db      121  TyrSerProLeuSerThrValProThrSerSerThrProProAlaLeuAspProSer  140
QY      3934  -----CCACAGTATGATGCTCGGAAAGGCTATATTTT  3969
Db      141  ProThrAlaSerSerSerProThrValSerAspAlaArgGlnGlySerIlePhe  160
QY      3970  TCTGTAGACACTAGTTTGGAGATCTCAGCGCTCAAAAAGGCTCATGTTCCGATAATAT  4029
Db      161  SerIleGluThrSerLeuGluIleSerGlyIleLysLysGluValMetPheAspAsn  180
QY      4030  GCGGGAATTCGGAACAGTTTTCGAGGTAAAGATAATAATCTGCTGGTGAGGC  4089
Db      181  AlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGlyGly  200
QY      4090  AGTGGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAAATAACTGTAAGGGAAGT  4149
Db      201  Ser-GlySerAlaThr-Thr-ProSer-----PheThrValLysAsnCysLysGlyLysVa  218
QY      4150  TTCCTTTCACAGATAAGTACGCTCTTCGAGGCGGAGTGCTTTATAAGGCAATTCGTCT  4209
Db      218  lSerPheThrAspAsnValAlaSerCysGlyGlyGlyValValTyrLysGlyThrValLe  238
QY      4210  TTTCAAAGCAATGAAGGAGGCATATTCTTCCAGGAGGAACACAGCATACGATCATTAAG  4269
Db      238  uphLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLeuGl  258
QY      4270  GATTTCTGCTGCTACTAATCAGGATCAG---ATACGGAGACAGGAGCGGCTGGAGGAGT  4326
Db      258  ylleLeuAlaAlaThr-SerGlnAspGlnSerAsnThrGluThrGlyGlyGlyGlyAl  278
QY      4327  TATTGCTTCCAGATGATCTGTAAGTTTGAAGCAATAAGGTTCTATTCGTTTGA  4386
Db      278  alleCysSerProGluAspThrValThrPheGluGlyAsnLysGlySerIleValPheAs  298
QY      4387  TTAACACTTTGCAAAAGCGAGCGGAGCATCTCAACGAAAGATTC  4435
Db      298  pTyrAsnPheAlaLysGlyArgGlyGlyIleLeuThrLysLysPhe  314

RESULT 16
PMP13 CHLPN STANDARD; PRT; 973 AA.
AC Q9Z896; Q9K2A1; Q9Z410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmp13 precursor (Polymorphic membrane
DE protein 13) (Outer membrane protein 14).
GN Name=pmp13; Synonyms=omp14;
OS OrderedLocustNames=CpN0453, Cp0299, CpB0470;
OC Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=92026606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).

```

```

RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Uterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwin M.L., Nelson W.C., DeBooy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Frazer C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-262.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AB001629; AAD18595.1; -; Genomic DNA.
CC EMBL; AB002191; AAF38156.1; ALT INIT; Genomic DNA.
CC EMBL; BA000008; BAA98660.1; -; Genomic DNA.
CC EMBL; AB017158; AAF98401.1; -; Genomic DNA.
CC EMBL; AJ133034; CAB37074.1; -; Genomic DNA.
CC PIR; B86547; B86547.
CC PIR; F72076; F72076.
CC PHCI-2DPAGE; Q9Z896; -.
CC TIGR; CP0299; -.
CC InterPro; IPR005546; Auto_transp_beta.
CC InterPro; IPR011427; ChlamPMP_M.
CC InterPro; IPR003368; Chlamydia_PMP.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; Chlam_PMP; 8.
CC Pfam; PF07548; ChlamPMP_M; 1.
CC TIGRFAMs; TIGR01376; POMP_repeat; 6.
CC Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
CC SIGNAL 1 24 Potential.
CC CHAIN 25 973 Probable outer membrane protein pmp13.
CC CONFLICT 258 258 N -> Y (in Ref. 5).
CC FT SEQUENCE 973 AA; 102762 MW; E02A69F61DEBFE2 CRC64;
CC
CC Alignment Scores: 7.27e-82 Length: 973
CC Pred. No.:

```



```
Db 1 GlySerIleIleSerAlaAsnGlyGlnAsnLeuThrIleValGlyGlnAsnHisThrLeu 80
Qy 3691 TCATTACAGATCTCAAGGCCAGTCTTCAAAATTAATGCTTCATTTACGACGAGAG 3750
Db 81 SerPheThrAspSerGlnGlyProAlaLeuGlnAsnCysAlaPheIleSerAlaGluGlu 100
Qy 3751 ACACCTACTCTGAGAGATTTTCGAGTCTGATGTTCTCGAAATATGTTCTTGGGAGAA 3810
Db 101 LysIleSerLeuArgAspPheSerSerLeuLeuPheSerLysAsnValSerCysGlyGlu 120
Qy 3811 AAGGAAATGATCTCCGGAAACCGTGAATTTCCGAGCAGCGGAGGAGTGAATTTCTCG 3870
Db 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyLysAspSerIleValPheLys 140
Qy 3871 GATAACTCCGTTGGGTATTTCTCTTTACTACTGTGCCAACCTCATCATCAACTCCGCT 3930
Db 141 AspAsnSerValGlyTyrSerSerLeu-----ProSerValGlyGlnThrProThr 157
Qy 3931 GCTCAACAGTAGTAGTCTCGGAAAGGCTATTTTCTGTGAGACACTAGTTTGGAG 3990
Db 158 ThrProIleValGlyAspValLeuLysGlySerIlePheCysValGluThrGlyLeuGlu 177
Qy 3991 ATCTCAGCGCTCAAAAGGCGTATGTTTCATATATATGCGCGGAATTTCCGAAACAGTT 4050
Db 178 IleSerGlyValLysLysGluLeuValPheAspAsnThrAlaGlyAsnPheGlyAlaVal 197
Qy 4051 TTTGAGGTAGAATAATAATATGCTGTGTGGAGCGAGTGGGTTCCGCTACACCATC 4110
Db 198 PheCysSer-Arg-----AlaAlaGlnGlu 205
Qy 4111 AAGTACGACTTTACAGTTAAACTGTAAAGGGAAGTCTTCTTTCACAGATACGTAGC 4170
Db 205 YAspThrThrPheThrValLysAspCysLysGlyLysIleLeuPheGlnAspAsnValGlu 225
Qy 4171 CTCTGCGGCGCGAGTGGTTTATAAGGATGTTGCTTTTCAAGACATGAAGGAGG 4230
Db 225 YSerCysGlyGlyValIleThrLysGlyGluValLeuPheGlnAspAsnGluGlyGlu 245
Qy 4231 CATATTTCTCGAGGAAACACAGCATACGATGATTTAAGGATTTCTGTCTACTATCA 4290
Db 245 uMetLeuPheArgGlyAsnSerAlaHisAspLeuGlyIleLeuAspAla---AsnPr 264
Qy 4291 GGATCAGATACGGACAGAGCGGTGGAGGATTTATTTGCTCTCAGATGATTTCTGT 4350
Db 264 oGlnProProThrGluValGlyGlyGlyValIleCysThrProGluLysThrVa 284
Qy 4351 AAAGTTTGAAGCAATAAAGTTCTATTGTTTGTGATTACAACTTTGCAGAAAGCAGAGG 4410
Db 284 lThrPheLysGlyAsnLysGlyProIleThrPheAspTyrAsnPheAlaLysGlyA-ggl 304
Qy 4411 CGAAGCATCTCAACGAAGAATTC 4435
Db 304 yGlyAlaIleGlnSerGlnThrPhe 312
```

```
RESULT 18
Q823X1_CHLVCV PRELIMINARY; PRT; 942 AA.
ID Q823X1 CHLVCV
AC Q823X1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family protein/autotransporter,
DE putative.
GN OrderedLocusNames=CCA00282;
OS Chlamydomophila caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83557;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
```

```
RA Heidelberg J.F., Holtzappple E.K., Khouri H.M., Pederoza N.B.,
RA Carly H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Heia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT *Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae.
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AF016995; AAP05033.1; -, Genomic_DNA.
DR TIGR; CCA00282; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto_transp_beta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
KW Complete proteome.
SQ SEQUENCE 942 AA; 100425 MW; E332BCB88507A912 CRC64;
```

Alignment Scores:

Pred. No.:	4.36e-68	Length:	942
Score:	1158.00	Matches:	326
Percent Similarity:	44.6%	Conservative:	153
Best Local Similarity:	30.4%	Mismatches:	402
Query Match:	14.7%	Indels:	192
DB:	2	Gaps:	31

US-10-701-844-1 (1-4435) x Q823X1_CHLVCV (1-942)

```
Qy 382 ATGCAACGCTTTCATTAAGTCTTCTTTCATTAAGTCTTAGCTTATCTTGTGCTCT 441
Db 1 MetLysHisProValTyrTrpPheLeuValSerSerGlyLeuLeuAlaSer---ThrSer 19
Qy 442 TTAATCGGGGGGATATCGACGAGAAATCATGTTCTTCAAGAAATTTACGATGGCGAG 501
Db 20 LeuSerPheAlaAlaThrValGlnGluThrLeuAsnSerSerAspSerTyrAsnGlyAsn 39
Qy 502 ACGTAACTCTGATCATTT-----CCCTATCTGTTTATA 534
Db 40 ThrAlaThrSerAlaPheGlnThrLysGluThrGlnAlaGlyAlaGluTyrThrCysGlu 59
Qy 535 GGAGATCCGAGTGGAGTACTGTTTCTGCGAGGAGTAAATTAATAAAATCTTGAC 594
Db 60 GlyAsn-----ValCysIleThrTyrAlaGlyLysGlyThr----- 71
Qy 595 AATTCATTGCGAGTTGCTTAAAGTGTGTTTGGGAATCTATTAGGAGTCTTACTGTT 654
Db 72 -----AlaLeuThrLysSerCysPheThrGluThrThrGluAsnLeuThrPhe 87
Qy 655 TTAGGAGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACAAATGGCGAGCT 714
Db 88 LeuGlyArgGlyTyrSerLeuCysPheAspAsnIleAsnThrThrAlaLysProAlaAla 107
Qy 715 CTAAGTAATAGCGTCTGATGAGTCTTTTATTTAGGGGTTTTAAAGAATTATCTTT 774
Db 108 IleGluValSerAlaAlaAspLysThrLeuSerIleSerGlyPheSerLeuPheSerCys 127
Qy 775 TCCAAATTGCAATTCACTTTCGCGTACTGCTGCAACGACTAATAAGGTTAGCCAG 834
Db 128 SerAspCys-----ProGlyThrThrGlyGlnGly----- 138
Qy 835 ACTCCGACGACACATCTACACCGTCTAATGTTACTATTATTCTTAAACAGATCTTTTG 894
Db 139 -----AlaIleLysSerGlyGlyThrAlaThrPheGluAsnAspPheSer 153
Qy 895 TTACTCAATTAATGAGAGTTCTCATTTATAGTAATTTAGTCTCTGAGATGGGGAGCT 954
Db 154 ValLeu-----PheLysLysAsnCysSerThrAlaAlaGlyGlyAla 167
Qy 955 ATAGATCCTAAGAGCTTAACGGTTCAAGGATTAAGCAAGCTTTGTGCTTCCAAAGAAAT 1014
```


Db 168 IleAsnCysLysGlyLeuThrLeuLysGlyThrSerGlyIleAlaAsnPhelIleGluAsn 187
Qy 1015 ACTGCTCAAGCTGATGAGGGGAGCTTGTCAAGTAGTACACAGTTTCTCGTACTATGCTAAC 1074
Db 188 LysSerThrAspAsnGlyGlyAla 195
Qy 1075 GAGGCTCTATTGCTCTTTGTACGAATGTTGCAGGAGTAAGAGGGGAGGATTGCTGCT 1134
Db 196 200
Qy 1135 GTTCAGGATGGCAGCAGGAGTGCATCATCTACTTCAACAGAAGATCCAGTAGTAAGT 1194
Db 201 212
Qy 1195 TTTTCCAGAAATCTGCGGTAGAGTTGTGAGGAAGTAGTACCGGAGTAGGAGGAGATT 1254
Db 213 PheSerGlyAsnThrSer 225
Qy 1255 TACTCTACGGGNAAGTTCCTTCTGTAATAATGGAATAACCTTGTCTCAACAATGTT 1314
Db 226 HisSerAsnSerAlaValThrIleAlaAsnAsnHisArgValGluPhe 241
Qy 1315 GCTTCTCTGTTTACATTGCTGCTAGCAACCAACAAAGTCGACAGGCTTCTAATACGAGT 1374
Db 242 253
Qy 1375 AATAATTACGAGATGAGGAGCTATCTTCTGTAAGAATGGTGGCAAGAGATCCAAT 1434
Db 254 262
Qy 1435 AACTCTGGATTCAGTTTCTTGTGAGGAGGAGTAGTTTCTTGTAGCAATGAGCT 1494
Db 263 AlaAlaProGluLeuLysPheGluGlyAsnAlaGlnLeuLeuPheLeuGluAsnSerSer 282
Qy 1495 GCTGGGAAGGGGAGCTATTATGCAAAAGCTCTCGTGTCTAACTGTGCGCCCTGTA 1554
Db 283 GlnValSerGlyGlyAlaLeuThrSerAsnLysLeuThrIleSerGlyGlyThrThr 302
Qy 1555 CAATTTTAAGGAATATCGCTAATGAT 1602
Db 303 ValPheAlaAsnAsnThrValThrAsnAlaAsnProMetGlyGlyAlaIleCysLeuAsp 322
Qy 1603 ---GAATCTGGAGAGCTCAGTTTATCTGCTGATTAAGAGATATTTTCGATGGGAAT 1659
Db 323 ThrThrSerGlyGlyCysSerLeuSerAlaAspLeuGlyAsnIlePheAspGlyAsn 342
Qy 1660 CTTAAAGACACCAAGAGATGCTGCGGATGTTAATGCGGTAACTGTGCTCCACAA 1719
Db 343 LysValIleThrSerGlyGlyAsn 356
Qy 1720 GCCATTTTCGATGGATCGGAGGGGAAATAACGACATTAAGAGCTAAGCAGGCGCATCAG 1779
Db 357 SerIleAspLeuAsnThrSerGlyLysPheThrGlnLeuArgAlaLysAspGlyPheGly 376
Qy 1780 ATTCTCTTTAATGATCCCATCGAGATGGCAAAACGGAATTAACCGCAGCGCAGTCTTCC 1839
Db 377 IleThrPheThrAspProIle 392
Qy 1840 AAATCTTAAATTAACGATGGTGAAGGA 1890
Db 393 ---LeuAsnIleAsnAlaProGluAsnAlaThrThrThrAsnGlyArgValAlaPhe 410
Qy 1891 GCT 1911
Db 411 SerGlyGluThrLeuSerAlaThrGluLysThrGluAlaAspAsnLeuLysSerIlePhe 430
Qy 1912 TACCAAAATGTTACGATAGACAAAGGAGGATGTTCTTCGTGAAAGGCAAAATTAATCA 1971
Db 431 LysGlnProValThrLeuSerAlaGlySerLeuIleLeuLysAspGlyValThrValGlu 450
Qy 1972 GTGAATCTCTAAGTCACAGAGTGGGAGT 2028
Db 451 AlaLysLysIleThrGlnThrAlaGlySerAlaValValMetAspAlaGlyThrThrLeu 470

Qy 2029 GATTTTGTAGTCTCCACCAACCAACAGAGCTCTCTGCCGCTAATATCATGATTGATCATCGCTT 2088
Db 471 482
Qy 2089 TCCAATCTGCATTTGTCTCTTTCTTTTGTAGCAACAATGCAATGCAATCTCTCTCT 2148
Db 483 ProAspLeuThrIleAsnValAlaSerPheGlyGlyGluGlyThr 498
Qy 2149 ACCAATCTCCAGCGCAAGATTCTCATCTGCAGTCATTGCTAGCACAACATGCTCGT 2205
Db 499 511
Qy 2206 TCTGTTACAAATTAGTGGCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGG 2265
Db 512 AlaLeuThrValThrAlaValSerPheIleAspAsp---AspGlyAsnGlyThrGluThr 530
Qy 2266 TATGATGGCTAGGTTCTAATCAAAAATCAATGCTCTGMAAATTAACAGTTAGGGATAG 2325
Db 531 ProValPheSerLysThrArgAspPheAlaAspSerIleLeuLeuGluAlaAlaThrGly 550
Qy 2326 CCCCAGCTAATGCCCATCA---GATTTGACTCTAGGAATGAGATGCTTAAGTATGCG 2382
Db 551 ThrThrValThrAlaProAlaIleProThrThrProAspThrProSerAlaHisThrGly 570
Qy 2383 TATCAAGGAAGCTGGAAGCTTGGTGGGATCCATAACAGCAAAATAATGGTCTTATATCT 2442
Db 571 TyrGlnGlyAsnThrThrIleAlaThrAlaGlnGlyThrAlaGlyThrHisGluGlnLys 590
Qy 2443 CTGAAGCTPACATGAGCTAAACTGGGTATTAATCTCGGCCCTGAGGAGTAGTCTTTTG 2502
Db 591 AlaThrLeuThrTrpThrGlyThrGlyThrIleProAsnProGluArgGlnAlaGlnLeu 610
Qy 2503 GTTCCAATAGTTTATGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTC 2562
Db 611 ValProAsnThrLeuTrpGlyAlaPheThrAspMetArgAlaLeuHisGlnLeuMetSer 630
Qy 2563 GCAAGTGTGATGGGCGCTCTATTGTCGAGGATTTGGTCTTCTGGAGTTTCGAATTC 2622
Db 631 ValSerAlaThrAspLeuGluGlnArgGlyLeuTrpGlyAlaAlaIleThrAspPhe 650
Qy 2623 TTCTATCATGACCGGATGCTTTAGTCAGGATATCGGTATATATAGTGGGTTTATCC 2682
Db 651 LeuGlnArgLysLysThrSerLysLysThrArgHisValGlyValGlyThrAla 670
Qy 2683 TTAGGACAACTCCCTACTTCTGGATCA---TCGATCTTGGTCTAGCATTTACCGAAGTA 2739
Db 671 ValGlyAlaSerValHisMetProThrGluAspLeuPheSerLeuAlaPheCysGlnPhe 690
Qy 2740 TTTGGTAGATCTAAAGATTTAGTGTGCTGTTCCAATCATCATCTGTCATAGGATCC 2799
Db 691 PheAsnAsnAspLysAspPheValValSerLysAsnArgThrHisValThrAlaGlySer 710
Qy 2800 GTTTAT---CTATCTACCAACAACAGCTTTTATGTTGATGCTTATTTG 2841
Db 711 LeuPhePheGluHisPheHisMetLeuHisProGlnAsnThrLeuLysValGlySerLys 730
Qy 2842 TTCGAGATGCGTTT 2862
Db 731 PheProAlaPheLeuAlaAsnLeuProGluAsnValProMetIleLeuAsnIleLeu 750
Qy 2863 GCTAGCTACGGTTCGGGAATCAGCATATGAAACCTCATATACA 2907
Db 751 PheSerThrSerHisAlaGluAsnAspMetLysThrArgThrLysArgThrSerPro 770
Qy 2908 ---TTTCAGAGGAGAGCGATGTTGCTGGGATTAATACTGCTGGCTGGAGAG 2958
Db 771 LysProValThrThrProGluValThrGlySerThrGlyThrAsnCysPheGlyGlyGlu 790
Qy 2959 ATTGGAGCGGATTCAGGATTTGATCTCCATCTAAGCTCTATTTGAATGAGTTGCT 3018
Db 791 IleSerThrSerPheProIleGluLeuSerAspSerThrMetPhe---GluArgPheVal 809

DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
 DE Polymorphic outer membrane protein.
 GN Name=pmp14g; Synonyms=pomp38A; OrderedLocusNames=CAB282;
 OS Chlamydomonas abortus.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiales; Chlamydiales.
 OX NCBI_TaxID=83555;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=S26/3;
 RX PubMed=15837807; DOI=10.1101/gr.3684805;
 RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
 RA Livingstone M., Cerdano-Tarrasa A.-M., Harris B., Doggett J.,
 RA Quail M.A., Price C., Clarke K., Fellwell T., Hance Z., Sanders M.,
 RA "The Chlamydomonas abortus genome sequence reveals an array of
 RT variable proteins that contribute to interspecies variation.";
 RL Genome Res. 15:629-640(2005).
 DR EMBL; CR848038; CAH63732.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 926 AA; 98440 MW; F81C2A79A3D575D1 CRC64;

 Alignment Scores:
 Pred. No.: 1,07e-65 Length: 926
 Score: 1122.00 Matches: 325
 Percent Similarity: 44.4% Conservative: 151
 Best Local Similarity: 30.3% Mismatches: 389
 Query Match: 14.2% Indels: 208
 DB: 2 Gaps: 35

 US-10-701-844-1 (1-4435) x QSL6J3_CHLAB (1-926)
 QY 382 ATGCAAGCTCTTCCATAGTCTTCTTCAATGATCTAGCTTATCTGCTGCTCT 441
 DB 1 MetArgProSerLeuThrLysLeuLeuLeuSerSerThrLeuThrPro---IleSer 19
 QY 442 TTAATGGGGGGGATATCGAGCAATCATGCTCTCTCAA---GGAAATTCAGATGG 498
 DB 20 PheHisPheSerGlnLeuHisAlaGluValAlaLeuThrGlnGluSerIleLeuAspAla 39
 QY 499 GAGAGCTTAACTGATCATCTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTT 558
 DB 40 Asn-----GlyAlaPheSerProGlnSerThr 48
 QY 559 TTTCTCGAGAGATTAACTAAATAAATCTTGCAATCTTAT----- 603
 DB 49 SerThrAlaGlyGly---ThrIleTyrAsnValGluSerAspIleSerIleValAspVal 67
 QY 604 -----GCAGCTTTCCTTAACTTGGTGGAACTTATAGGAGTTTACTGTT 654
 DB 68 GlyGlnThrAlaAlaLeuAlaSerSerAlaPheValGlnThrAlaAspAsnLeuThrPhe 87
 QY 655 TTAGGAGAGACACTCTGACTTTCGAGACATACGACTTCTACAAATGGGCGAGCT 714
 DB 88 LysGlyAsnAsnHisSerLeuSerIleThrAsnAlaAsnAlaGlyAlaAsnProAlaGly 107
 QY 715 CTAAGTAATAGCGCTGATGAGCTGTTTACTATTTAGGGTTTAAAGATTATCTTT 774
 DB 108 IleAsnValAsnThrAlaAspLysIleLeuThrLeuThrAspPheSerLysLeuSerPhe 127
 QY 775 TCCAATTCGCAATTCATTCTGCTGCTGCTGCAACGACTTAATAAGGTTAGCCAG 834
 DB 128 LysGluCys-----ProSerSerLeuValAsnThrGly----- 138
 QY 835 ACTCCGACGACCAACATCTACCGCTCTAATGGTACTATTATTTCTAAACAGATCTTTG 894
 DB 139 -----LysGlyAlaMetLysSerGlyAlaLeuAsn 149
 QY 895 TTACTCAATATAGAGATTCTCATCTTATAGTAATTTAGTCTCGAGATGGGGAGCT 954
 DB 150 LeuAlaAsnAsnAlaSerIleLeuPheAspGlnAsnTyrSerAlaGluAsnGlyGlyAla 169
 QY 955 ATAGATCTAGAGCTTAAACGGTTCAAGGAATTAGCAAGCTTGTGCTCTCCAGAAAT 1014

Db 170 IleSerCysLysAlaPheSerLeuThrGlySerSerLysGluIleSerPheThrAsn 189
 QY 1015 ACTGCTCAAGCTGATGGGAGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAAC 1074
 Db 190 SerThrAlaLysLysGlyGlyAla----- 197
 QY 1075 GAGGCTCTATTGCTCTTTAGCGAATGTTGCAGGAGTAAGAGGGGGGAGGATTCGCTCT 1134
 Db 198 -----IleAlaAlaThrGlyIleAlaHis 205
 QY 1135 GTTCAGGATGGGAGGAGGAGTGTCTATCTACTTCAACAGAGATCCAGTAGTAGT 1194
 Db 206 LeuSerAspAsnGln-----GlyThrIleArg 214
 QY 1195 TTTTCCAGAATACTCGGTAGAGTTTGTATGGGAACGTAGCCGAGTAGGAGGAGGAT 1254
 Db 215 PheSerGlyAsnThrAlaVal-----AsnSerGlyGlyAlaVal 227
 QY 1255 TACTCTACGGGAACGTGCTTCTCTCAATAATGAAAAACCTTGTCTCTCAACATGTT 1314
 Db 228 TyrSerGluAlaSerMetThrIleAlaGlyAsnAsnHisValAlaPheSerAsnAla 247
 QY 1315 GCTTCTCTGTTTACATTGCTGTGCTAAGCAACCAAGTGGACAGGCTTCTAATACGAGT 1374
 Db 248 ValSer-----Gly 250
 QY 1375 AATAATTACGAGATGGAGGAGCTATCTTCTGTAAGAATGGTGCAGCAGGAGTCCAA 1434
 Db 251 SerSerAspGlyCysGlyGlyAlaIleHisCysSerLys-----ThrGly 265
 QY 1435 AACTCTGATCAGTTCTCTTGTATGGAGAGGAGTAGTATTTCTTGTAGTAGCAATGTAGT 1494
 Db 266 SerAlaProThrLeuThrIleArgAspAsnLysValLeuIlePheGluGluAsnThrSer 285
 QY 1495 GCTGGGAAGGGGAGCTATTATGCCAAAAGCTCGGTGCTTAATCTGTCGCCCTGTA 1554
 Db 286 SerAlaLysGlyAlaIleThrAspLysLeuLeuThrSerGlyGlyProThr 305
 QY 1555 CAATTTTAAAGGAT-----ATCGCTAATGATGTTGAGCGAGTTTATTTAGGA 1602
 Db 306 AlaPheIleAsnLysValThrHisAlaThrProLysGlyGlyAlaIleGlyIleAla 325
 QY 1603 GAATCTGAGAGCTCAGTTATCTGCTGATATGAGATATTTATTTTCGATGGGAATCTT 1662
 Db 326 AlaAsnGlyGluCysSerLeuThrAlaGluHisGlyAspIleThrPheAspAsnAsnLeu 345
 QY 1663 AAAAGAACAGCAAGAGATGCTGCCGATGTTAATGCGGTAACTGTCCTCACAAGCC 1722
 Db 346 MetAlaThr---GlnAspAsnAla-----ThrIleLysArgAsnAla 358
 QY 1723 ATTTTCGATGGGATCGGAGGAGGAAAATAACGACATTAAGAGCTTAAGCAGGCGCATCAG 1782
 Db 359 IleAsnIleGluGlyAsnGlyLysPheValAsnLeuArgAlaAlaSerGlyLysThrIle 378
 QY 1783 CTCTTTAATGATCCATCGAGATGGCAACGGAAATAACAGCCAGCGCGCTCTTCCAAA 1842
 Db 379 SerPheTyrAspProIleThrVal---GluGlyAsn-----AlaAlaAsp 392
 QY 1843 CTCTTAAATAATACGATGGTGAAGGA-----TACACAGGGGATATTGTTTTGCT 1893
 Db 393 LeuLeuThrLeuAsnLysAlaGluGlyAspLysThrTyrAsnGlyArgIleIlePheSer 412
 QY 1894 -----AATGGAAGCGTACTTGTGTAC 1914
 Db 413 GlyGluLysLeuThrGluGluAlaAlaValAlaAspAsnLeuLysThrThrPheThr 432
 QY 1915 CAAAATGTTACGATAGCAGCAGGAGGATTTGTTCTTCTGTAAGGCAAAATTTATCATGTG 1974
 Db 433 GlnProIleThrLeuAlaAlaGlyValLeuValLeuArgSerGlyValGluValGluAla 452
 QY 1975 AATTCTCTAAGTACAGAGGTTGGGAGTCTG---TATATGGAAGCTGGGAGTACATGGAT 2031
 Db 453 LysThrValValGlnThrAlaGlySerLeuIleLeuMetAspAlaGlyThr----- 469

QY 2032 TTTGTAACCTCCAAACCAACACACAGCCTCTCGCGCTTAATCACTGATCAGCGCTTCC 2091
DB 469 ----- 469
QY 2092 AATCTGCATTTGCTCTTTCTTTCTTTGTAGCAAAATGCAATGCAATCTCCTACC 2151
DB 470 -----LysLeuSerAlaLysThrGluAspAlaThrLeuThrAsnLeuAlaIle 485
QY 2152 AATCTCCAGCCAGATCTCATCTCT---GCAGTCATTGGTAGCACTCTCTGGT--- 2205
DB 486 AsnProAsnThrLeuAspGlyLysLysPheAlaValAlaAspAlaValAlaAlaGlyLys 505
QY 2206 TCTGTACATTAAGTGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGG 2265
DB 506 AsnValThrLeuSerGlyAlaIleGlyValIleAspProThrGlyLysPheTyrGlu--- 524
QY 2266 TATGATTGGCTAGGTTCTAATCAAAAATCAAT-----GTCTGAAATATACAG 2313
DB 525 -----AsnHisLysLeuAsnAspThrLeuAlaLeuGlyIleGln 538
QY 2314 TTA-----GGGACTAAGCCCCAGCTAATGCCCATCAGATTGACTTAGGGAAT 2364
DB 539 LeuSerGlyLysGlySerValThrThrAsnValProSerHisVal---ValGlyVal 557
QY 2365 GAGATCCCTAAGTATGCTATCAGGAAGCTGGAAGCTTGGTGGATCCTAATACAGCA 2424
DB 558 AlaGluThrHisTyrGlyTyrGlnGlyAsnTrpSerValSerTrpValLysAspAsn--- 576
QY 2425 AATAATGGTCTTATCTCTGAAAGCT-----ACATGACTAAAACTGGGTATATCTCT 2478
DB 577 AsnSerAspProLysThrGlnThrAlaIlePheThrTrpAsnLysThrGlyTyrValPro 596
QY 2479 GGGCTGAGCGAGTACTCTTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
DB 597 AsnProGluArgAlaProLeuValLeuAsnSerLeuTrpGlySerPheIleAspLeu 616
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGG---CGCTCTTATTGTCAGCA 2595
DB 617 ArgSerIleGlnAspValLeuArgSerValAspSerIleLeuGluThrArgArgGly 636
QY 2596 TTATGGGTTTCTGGAGTTCTGAAATCTCTATCATGACCGCATGCTTTAGTCAAGCA 2655
DB 637 LeuTrpValSerGlyIleGlyAsnPheHisLysAspArgAsnAlaGluAsnArgLys 656
QY 2656 TATCGGTATATTAGTGGGGTTATCTCTTAGGAGCAAACTCTCTACTTTGCA---TCATCG 2712
DB 657 PheArgHisIleSerSerGlyTyrValLeuGlyAlaThrThrAsnThrSerArgGluAsp 676
QY 2713 ATGTTTGGTCTAGCATTTACGAGATATTGGTAGATCTTAAAGATTATGATGTCGT 2772
DB 677 SerLeuSerValAlaPheCysGlnLeuPheAlaLysAspLysAspTyrLeuValSerLys 696
QY 2773 TCCATCATCATCTTGCATAGGATCGGTTAT----- 2805
DB 697 AsnAlaAlaAsnValTyrAlaGlySerValTyrTyrGlnHisValSerLysPheAspAsp 716
QY 2806 -----CTATCTACCCCAACAGCTTTATGTGGATCCCTAT----- 2838
DB 717 LeuThrArgLeuPheAsnGlyProAsnThrCysCysSerGlyPheSerLysGluIlePro 736
QY 2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAATAACC 2898
DB 737 IlePheLeuAspAlaGlnIle-----ThrTyrCysHisThrAlaAsnAsnMetThrThr 754
QY 2899 TCATATACATTTCAGAGGAGGAGCGATGTCGTGGGATATAACTCTCGGCTGGGAGAG 2958
DB 755 SerTyrThrAspTyrProGluValLysGlySerTrpGlyAsnAspThrLeuGlyLeuThr 774
QY 2959 ATTGGAGCGGATPACCGATTGTTGATTACTCCATCTAAGTCTATTGTAAGTGTTCGT 3018
DB 775 LeuSerThrSerValProIleProValPheSerSerSerIlePhe---AspSerTyrAla 793

QY 3019 CCTTTGCGCAAGCTGAGTTTCTTTATGCCCATCATGAATCTTTTACAGAGAGGCGCAT 3078
DB 794 ProPheAlaLysLeuGlnValTyrAlaHisGlnAspAspPheLysGluProThrThr 813
QY 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTTAATCTATCATGCTCTCTGTTGGAGTGAAG 3138
DB 814 GluGlyArgValPheGluSerSerAspLeuLeuAsnValSerValProIleGlyIleLys 833
QY 3139 TTTGATCGCATGTTCTAGTACACATCTTAATAATATAGCTTTTATGGCGGCTTATCTGT 3198
DB 834 PheGluLysLeuSerTyrGlyGluArgSerAlaTyrAspLeuThrLeuMetTyrIlePro 853
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAGCTCTCTATCCATCAGAGACATGG 3258
DB 854 AspValTyrArgHisAsnProSerCysMetThrGlyLeuAlaIleAsnAspValSerTrp 873
QY 3259 ACAACAGATGCTTCTTATTTAGCAAGACATGGAGTTGTTGGTAGAGGATCTATGTATGCT 3318
DB 874 LeuThrThrAlaThrAsnLeuAlaArgGlnAlaPheIleValArgAlaGlyAsnHisIle 893
QY 3319 TCTCTAAACAAGTAATATAAGATATATGCCCATGGAAGATATGATCGAGATCTTCT 3378
DB 894 AlaLeuThrSerGlyValGluMetPheSerGlnPheGlyPheGluLeuArgSerSerSer 913
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
DB 914 ArgAsnTyrAsnValAspLeuGlyAlaLysValAlaPhe 926
RESULT 21
PMP10 CHLPN STANDARD; PRT; 928 AA.
AC Q9RB65; O86163; Q9RB64; Q9S6P2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp10 precursor (Polymorphic membrane
protein 10) (Outer membrane protein 5).
GN Name=pmp10; Synonyms=omp5; OrderedLocusNames=CP0303, CpB0467;
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=833558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christensen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
Madisen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
role in immunopathogenicity."
RL Am. Heart J. 138:S491-S495(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
Bass S., Linher K.D., Weidman J.P., Khouri H.M., Craven B., Bowman C.,
Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN-TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensach K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN (5)
RP NUCLEOTIDE SEQUENCE OF 1-914.
RC STRAIN-CW029/VR-1310;
RX MBLINS-99081766; PubMed:9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer
RT membrane proteins of Chlamydia pneumoniae.";
RL Infect. Immun. 67:375-383(1999).
CC -! SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -! SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AJ133034; CAB37071.1; -; Genomic DNA.
DR EMBL; AB021292; AAF38160.1; -; Genomic DNA.
DR EMBL; BA000008; BAA98657.1; -; Genomic DNA.
DR EMBL; AB017158; AAF98398.1; -; Genomic DNA.
DR EMBL; AJ001311; CAA04671.1; -; Genomic DNA.
DR PIR; G81591; G81591.
DR PIR; G86546; G86546.
DR PHCI-2DPAGE; O86163; -.
DR TIGR; CP0303; -.
DR InterPro; IPR005546; Auto transportbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR006826; PPH1.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam PMP; 7.
DR Pfam; PF07548; ChlamPMP M; 1.
DR SMART; SM00710; PPH1; 4.
DR TIGRPFAMs; TIGR01414; autotrans bar1; 1.
DR TIGRPFAMs; TIGR01376; POMP_repeat; 5.
KW Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 928 Probable outer membrane protein pmp10.
FT CONFLICT 305 305 Missing (in Ref. 4).
SQ SEQUENCE 928 AA; 97230 MW; 0590D5206A1DD0E1 CRC64;

Alignment Scores:
Pred. No.: 2,65e-64 Length: 928
Score: 1101.00 Matches: 324
Percent Similarity: 44.5% Conservative: 150
Best Local Similarity: 30.5% Mismatches: 402
Query Match: 14.0% Indels: 188
DB: 1 Gaps: 34

US-10-701-844-1 (1-4435) x PMP10_CHLPP (1-928)
QY 382 ATGCAACGCTTTCATTAAGTCTTTCTTCAATGATTCTAGCT--TATTCCTGCTGC 438
DB 1 MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys 20
QY 439 TCTTTAAATGGGGGGGATATGCACGAAATCATGGTCTCTCAAGGAATTTACAGTGG 498
DB 21 SerThrValPheAlaAla---ThrAlaGluAsnLeuGlyProSerAspSerPheAspGly 39
QY 499 GAGACGTTAAGTATCATTTCCCTATATCTGTATAGAGATCCGAGTGGGACTCTGTT 558
DB 40 SerThrAsnThrGlyThr-----TyrThrProLysAsnThrThrThrGlyIleAspTyr 57
QY 559 TTTTCTGCAGGAGAGTTAACAATTAATAAAATCTTGACATTCATTTGCAGCTTTGCTTTA 618

Db 58 ThrLeuThrGlyAspIleThrLeuGluAsnLeuGlyAspSer---AlaAlaLeuThrLys 76
QY 619 AGTTGTTTGGGAACATTTATTAGGAGGTTTACTGTTTATGGGAGAGGACACTCGTGTACT 678
DB 77 GlyCysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSer 96
QY 679 TTCGAGAACATACGGACTTCTACAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGGA 738
DB 97 PheLeuAsnIleLysSerSerAlaGluGlyAlaAlaLeuSer---ValThrThrAspLys 115
QY 739 CTGTTTACTATGAGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTTACTTGC 798
DB 116 AsnLeuSerLeuThrGlyPheSerSerLeuThrPhe-----LeuAlaAla 130
QY 799 GTACTGCTCTGCACAGCACTAATAGGGTAGCCAGACTCCGACGACACATCTACACCG 858
DB 131 ProSerValIleThrThrProSerGly----- 140
QY 859 TCTAATGCTACTATTATTCTTAAACAGATCTTTTGTGTACTCAATATGAGAATCTCTCA 918
DB 141 ---LysGlyAlaValLysCysGlyGlyAspLeuThrPheAspAsnAsnGlyThrIleLeu 159
QY 919 TTCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACCGTT 978
DB 160 PheLysGlnAspTyrCysGluGluAsnGlyGlyAlaIleSerThrLysAsnLeuSerLeu 179
QY 979 CAAGGAATTAGCAAGCTTTGTGTTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCT 1038
DB 180 LysAsnSerThrGlySerIleSerPheGluGlyAsnLysSerSerAla----- 195
QY 1039 TGTCAGTAGTCACCACTTCTCTGCTATGCTTAACGAGCTCTATTGCTTTGTAGCG 1098
DB 195 ----- 195
QY 1099 AATGTTGCAGGATTAAGAGGGGAGGATTGCTGCTGTTTCAGGATGGCAGGAGGTG 1158
DB 196 -----ThrGlyLysLysGlyGlyAlaIleCysAla----- 205
QY 1159 TCATCATCTACTTCAACAGCAAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATACGTGCG--GTA 1215
DB 206 -----ThrGlyThr-----ValAspIleThrAsnAsnThrAlaProThr 218
QY 1216 GAGTTTCAATGGAACTAGCCCGA---GTAGGAGGAGGATTTACTCTCAGGGGAAAGCTT 1272
DB 219 LeuPheSerAsnAsnIleAlaGluAlaAlaGlyGlyAlaIleAsnSerThrGlyAsnCys 238
QY 1273 GCTTTCCTGAATTAATGAAAACCTGTTTCTCAACAAATGTTGCTTCTCCTGTTTACATT 1332
DB 239 ThrIleThrGlyAsnThrSerLeuValPheSerGluAsn----- 251
QY 1333 GCTGCTAAGCAACCAACAGTGGACAGCTTCTAATACGAGTAATAATTACGAGATGGA 1392
DB 252 -----SerValThrAlaThrAlaGlyAsnGly 260
QY 1393 GGAGCTATCTTCTGAAGATGGTGGCCAGCAGGATCCCAATTAATCTGATCATGTTTCC 1452
DB 261 GlyAlaLeu-----SerGlyAspAlaAspValThr 270
QY 1453 TTTGATGAGAGGAGTAGTTTCTTTTAGTAGCAATGCTAGCTGCTGGGAAAGGGGAGCT 1512
DB 271 IleSerGlyAsnGlnSerValThrPheSerGlyAsnGlnAlaValAlaAsnGlyGlyAla 290
QY 1513 ATTTATGCCAAAAGCTCTCGTTTGTCTAAC-----TGTGGCCTGTACAAATTTTAAAG 1566
DB 291 IleTyrAlaLysLysLeuThrLeuAlaSerGlyGlyGlyIleSerPheSerAsn 310
QY 1567 AATATC-----GCTAATGATGGTGGAGCGGATTATTATTAGGAGAACTCGGA 1611
DB 311 AsnIleValGlnGlyThrThrAlaGlyAsnGlyGlyAlaIleSerIleLeuAlaAlaGly 330
QY 1612 GAGCTCAGTTTATCTGCTGATTTATGGAGATATTATTTCCATGGGAAT----- 1659

331 GluCysSerLeuSerAlaGluAlaGlyAspIleThrPheAsnGlyAsnAlaIleValAla 350
1660 ----CTTAAAGAACCCAAAGAGATGTCGCGATGTTAATGGCGTAATCTGTCTCA 1716
351 ThrThrProGlnThrThrLysArgAsnSerIleAspIle----- 363
1717 CAAGCCATTTCGATGGATCGGAGGCGAAATAACGACATTAAAGACCTAAAGCAGGCGAT 1776
364 -----GlySerThrAlaLysIleThrAsnLeuArgAlaIleSerGlyHis 378
1777 CAGATTCTCTTAAATGATCCCATCGAGATGGCAACGAAATACCGACGAGCGCATCT 1836
379 SerIlePhePheThrAspProIleThrAla-----AsnThrAlaAlaAspSer 394
1837 TCCAAACTCTAAAATTACGATGCTGAAGG-----TACACAGGGGAT 1881
395 ThrAspThrLeuAsnLeuAsnLysAlaAspAlaGlyAsnSerThrAspThrAspThrGlySer 414
1882 ATTGTTTCTGCT-----AATGGAAGC 1902
415 IleValPheSerGlyGluLysLeuSerGluAspGluAlaLysValAlaAspAsnLeuThr 434
1903 AGTACTTTGACCAAAATGTTACGATAGACAAAGGAGATTGTTCTCGTGAAGGCA 1962
435 SerThrLeuLysGlnProValThrLeuThrAlaGlyAsnLeuValLeuLysArgGlyVal 454
1963 AAATTATCAGTGAATCTCTAAGTCACACA--GGTGGAGTCTGTATATGGAAGCTGGG 2019
455 ThrLeuAspThrLysGlyPheThrGlnThrAlaGlySerSerValIleMetAspAlaGly 474
2020 AGTACATGGGATTTGTTAACTCCACACCAACACAGCCTCTCGCGCTAATCAGTTG 2079
475 ThrThrLeu-----LysAlaSerThrGluGlu 483
2080 ATCAGCTTTCCTCAATCTGCTTTCTTTCTTTCTTTGTTAGCAACAATGCGATTACG 2139
484 ValThrLeuThrGlyLeuSerIleProValAspSerLeu----- 496
2140 AATCCTCTACCAATCTCCAGCGCAAGATCTCATCTCGAGTCATCTGTCAGCAACT 2199
497 -----GlyGluGlyLysLysValValIleAlaAlaSerAlaAla 509
2200 GCTGGTCTCTTACAAATGATGGCGCTATCTTTTTCGAGGATTTGGATGATACAGCTTAT 2259
510 SerLysValAlaLeuSerGlyProIleLeuLeuLeuAspAsnGlnGlyAsnAlaLys 529
2260 GATAGGTATGATGGCTAGTCTTAATCAAAAAATCAATGCTCTGAAATACAG----TTA 2316
530 GluAsnHisAsp--LeuGlyLysThrGlnAspPheSerPheValGlnLeuSerAlaLeu 548
2317 GGGACTAAGCCCCCAGCTAATGCCCATCAGATTGATCTAGGGAATGAGTGCCTAAG 2376
549 GlyThrAlaThrThrAspValProAlaValProThrValAlaThrPro--ThrHis 567
2377 TATGGCTATCAGGAAGCTGAAGCTGGCGGATCCTTAATACAGCAAAATAATGTCCT 2436
568 TyrGlyThrGlnGlyThrThrPglyMetThrThrValAspAspThrAlaSerThr--Pro 586
2437 TATACTCTGAAAGCTACA-----TGGACTAAAACCTGGGTATATCTCGGCGCTGAGCGA 2490
587 LysThrLysThrAlaThrLeuAlaThrThrAsnThrGlyThrLeuProAsnProGluArg 606
2491 GTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATATACGATCTCGCGAT 2550
607 GlnGlyProLeuValProAsnSerLeuTrpGlySerPheSerAspIleGlnAlaIleGln 626
2551 TCAGCAATCAAGCAAGTGTGGTGGGCTCTTATTGTCAGGATATTGGTTCTTGGGA 2610
627 GlyValIleGluArgSerAlaLeuThrLeuLysSerAspArgGlyPheThrAlaAlaGly 646
2611 GTTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGCTCAGGATATCGGTATATTAGT 2670
647 ValAlaAsnPheLeuAspLysLysGlyGluLysArgLysThrArgHisLysSer 666

2671 GGGGTTTATTCCTTA---GGAGCAAACTCTACTTTGGATCATCGATGTTGGTCTAGCA 2727
667 GlyGlyThrAlaIleGlyAlaAlaGlnThrCysSerGluAsnLeuIleSerPheAla 686
2728 TTTACCGAAGTATTTGGTAGATCTAAAGATATATGATGTGTCTGTTCCATCATCATGCT 2787
687 PheCysGlnLeuPheGlySerAspLysAspPheLeuValAlaLysAsnHisThrAspThr 706
2788 TGCATAGGATCGTTTATCTATCT-----ACC 2814
707 TyrAlaGlyAlaPheThrIleGlnHisIleThrGluCysSerGlyPheIleGlyCysLeu 726
2815 CAACAAGCTTATGTGGATCTTATTTGTTGGAGATGCGTTTATC-----CGTCTACG 2868
727 LeuAspLysLeuProGlySerTrpSerHisLysProLeuValLeuGluGlyGlnLeuAla 746
2869 TACGGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGATGTT 2928
747 TyrSerHisValSerAsnAspLeuLysThrLysThrAlaLysProGluValLysGly 766
2929 CGTTGGGATAATAACTGCTCTGCTGGAGAGATTGGAGCGGATACCGATTGTGATTACT 2988
767 SerTrpGlyAsnAsnAlaPheAsnMetMetLeuGlyAlaSerSerHisSerTyr----- 784
2989 CCATCTAAGCTCTAT---TTGAATGAGTTGCGTCTCTTTCGTCACAGCTGAGTTTCTTAT 3045
785 ProGluThrLeuHisCysPheAspThrTyrAlaProTyrIleLysLeuAsnLeuThrTyr 804
3046 GCGGATCATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCGATTCAGAGCGGACAT 3105
805 IleArgGlnAspSerPheSerGluLysGlyThrGluGlyArgSerPheAspAspSerAsn 824
3106 CTCCTAAATCTATCAGTCTCTGTTGGAGGAAATTTGATCGATGTTCTAGTACACATCCT 3165
825 LeuPheAsnLeuSerLeuProIleGlyValLysPheGluLysPheSerAspCysAsnAsp 844
3166 AATAAATATAGCTTTATGCGGCTTATCTGTGATGCTTATCGCACCATCTCTGCTGACT 3225
845 PheSerTyrAspLeuThrLeuSerTyrValProAspLeuIleArgAsnAspProLysCys 864
3226 GAGACAACGCTCTATCCCATCAAGACAGATGAGCAACAGATGCTCTTCTTTAGCAAGA 3285
865 ThrThrAlaLeuValIleSerGlyAlaSerTrpGluThrTyrAlaAsnAsnLeuAlaArg 884
3286 CATGGATGTGCTTGAAGGATCTATGATGCTTCTCTCAACAGTAATATAGAAGTATAT 3345
885 GlnAlaLeuGlnValArgAlaGlySerHisTyrAlaPheSerProMetPheGluValLeu 904
3346 GGCCATGGAAGATATCAGTATCGAGATGCTTCTCGAGGCTATGGTTTGTAGTCAGGAAGT 3405
905 GlyGlnPheValPheGluValArgGlySerSerArgIleTyrAsnValAspLeuGlyGly 924
3406 AGAGTCCGGTTC 3417
925 LysPheGlnPhe 928

RESULT 22

Q823W9 CHLCV
ID Q823W9 CHLCV PRELIMINARY; PRT; 926 AA.
AC Q823W9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family
DE protein/autotransporter.
DE OrderedLocusNames=CCA00284;
GN Chlamydophila caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_Taxid=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;

```

RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapfle E.K., Khouri H.M., Fedorova N.B.,
RA Carthy H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Haia R.-C., McClarty G., Rank R.G.,
RA Bavail P.M., Fraser C.M.;
RT "Genome sequence of Chlamydomophila caviae (Chlamydia peittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147 (2003).
DR EMBL; AE016995; AAP05035.1; -; Genomic_DNA.
DR TIGR; CCA00284; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Auto_transportheta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR SMART; SMO0710; Pbh1; 4.
DR TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 5.
KW Complete proteome.
SQ SEQUENCE 926 AA; 98213 MW; 0E1062E0778658B2 CRC64;

Alignment Scores:
Pred. No.: 3,6e-64 Length: 926
Score: 1099.00 Matches: 315
Percent Similarity: 43.8% Conservative: 154
Best Local Similarity: 29.4% Mismatches: 398
Query Match: 13.9% Indels: 204
DB: 2 Gaps: 33

US-10-701-844-1 (1-4435) x QB23W9_CHLCV (1-926)
QY 382 ATGCAACGCTTCCCAATAGTTCTTTCTTCAATGATCTAGCTTATCTTCTGCTGCT 441
DB 1 MetArgProSerLeuTyrLysIleLeuIleSerSerThrLeuThrIlePro---LeuSer 19
QY 442 TTAATGGGGGGGATATGCGACAGAAATATGTTCTCTCAA----- 483
DB 20 PheHisPheSerGlnValHisAlaGluValAlaLeuThrGlnGluSerValLeuAspAla 39
QY 484 ---GGAAATTTACGATGGGAGACGTAACTGTATCAATTTCCCTACTATGTTATAGGAGAT 540
DB 40 AsnGlyAlaPheSerProGlnSerThrSerThrAla----- 51
QY 541 CGAGTGGGACTACTGTTTTTCTGCGAGGAGTTAACTAAATTAATAATCTTGACAAATCT 600
DB 52 ---GlyGlyThrThrTyrAsnValGluSerAspIleSerIleValAspAlaGlyGlnAla 70
QY 601 ATTGACGCTTGCCTTAAAGTTGTTTGGGAACCTATTAGGAGTTTACTGTTTAGGG 660
DB 71 ---AlaAlaMetValSerAlaAlaPheValGlnThrAlaAspLeuThrPheLysGly 89
QY 661 AGAGACACTCGTTGACTTTCAGAACATACGAGCTTCTACAAATGGGGGAGCTCTTAAGT 720
DB 90 AsnGlyAgsSerLeuAlaIleGluAsnValAsnSerGlyAlaAsnProGlyAlaIleTyr 109
QY 721 AATAGCGCTGCTGATGGACTGTTTACTATTAGGGTTTAAAGAAATATCTTTTCCAAT 780
DB 110 ValSerAlaAlaAspLysThrLeuThrLeuThrAspPheSerThrLeuSerPheLysLys 129
QY 781 TGCAATTCATTCTTCCGCTACTGCTGCTCAACGACTAATAAGGTTAGCGAGCTCGG 840
DB 130 Cys-----ProLysHisThrValAsnThrGly----- 138
QY 841 ACGAACACATCTACACCGCTCTAATGCTACTATTATTTCTTAAACAGATCTTTTGTATCTC 900
DB 139 -----LysGlyAlaValLysSerGlyGlyAlaLeuAsnLeuAla 151

```

```

QY 901 AATAATGAGAGTCTTCATCTTATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGAT 960
DB 152 AsnAsnAlaSerIleLeuPheAsnGlnAsnHisSerAlaGluAspGlyGlyAlaIleSer 171
QY 961 GCTAAGAGCTTAAACGGTTCAAGGAATTAAGCAAGCTTTGTGTCTTCCCAAGAAATACTGCT 1020
DB 172 CysLysAlaPheSerLeuThrGlySerLys----- 182
QY 1021 CAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAACGAGGCT 1080
DB 182 ----- 182
QY 1081 CCTATTGCCTTTAGCGAATTTGCAGAGTAGAAGAGGGGAGGATTCGCTGTTTCAG 1140
DB 183 GluIleSerPheThrThrAsnThrSerThrLysGlyGlyAlaIleAlaAlaThr--- 201
QY 1141 GATGGCGAGCAGGAGGTGTCATCTTCAACAGAAGATCCAGTAGTAGTAAAGTTTTC 1200
DB 202 -----GlyValAlaAsnLeuSerAspAsnGlnGlyLysValIle---PheSer 216
QY 1201 AGAAATACGCGTAGAGTTTGTATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTTC 1260
DB 217 GlyAsnThrAlaVal-----AenSerGlyGlyAlaValTyrAla 229
QY 1261 TAGGGGAACGTTGCTTCTCGAATAATAGGAACCTGTTTCTCAACAATGTTGCTTCT 1320
DB 230 GluAlaAsnThrThrIleAlaGlyAsnSerAlaValValPheAsnAsnAlaVal--- 248
QY 1321 CCTGTTTACATTTGCTGCTAAGCAACCAAGTGGACAGGCTTCTAATACAGTAATAAT 1380
DB 249 -----ThrGlyThrThr 252
QY 1381 TAGGAGATGGAGGAGCTATCTTCTGT---AAGAAATGTCGCGCAAGCAGGATCCAATAAC 1437
DB 253 AspGlyCysGlyGlyAlaIleHisCysSerLysThrGlyAla----- 266
QY 1438 TCTGATCAGTTTCTTTGATGGAGGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCT 1497
DB 267 ThrProValLeuThrIleArgAspAsnLysValLeuLeuPheLysGluAsnThrSerAla 286
QY 1498 GGGAAAGGGGAGCTATTATGCCAAAAGCTTCGCTGCTGCTAACTGTCGCCCTGTACAA 1557
DB 287 AlalysGlyGlyAlaIleTyrAlaAspLysLeuTyrLeuThrSerGlyGlyProThrVal 306
QY 1558 TTTTAAAGGAATATCGCTAATAGAT-----GGTGGAGCGATTTATTTAGGAGAA 1605
DB 307 PheValGlyAsnLysAlaThrAsnAlaAlaProLysGlyGlyAlaIleGlyIleAlaAla 326
QY 1606 TCTGAGAGCTCAGTTTATCTGCTGATATGAGATATATTTCGATGGGAATCTTAA 1665
DB 327 AsnGlyGluCysSerIleThrAlaGluHisGlyAspIleThrPheGluAsnAsnLeuIle 346
QY 1666 AGAACGCCAACAGAGATGCTCGCGATTAATGGCGTAACTGTGCTCTCACAGCCAT 1725
DB 347 AlaThrAla-----AsnAsnAlaThrValLysArgAsnAlaIle 359
QY 1726 TCGATGGGATGGGAGGAGAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTC 1785
DB 360 AsnIleGluGlyAsnGlyLysPheValAsnLeuArgAlaAlaSerGlyAsnThrLeuThr 379
QY 1786 TTTAATGATCCATCGAGATGGCAACGGAATAACAGCCAGCGGAGCTTCCAAACTT 1845
DB 380 PheTyrAspProIleValValGly-----GlyThrAlaAlaAspLeu 393
QY 1846 CTAAAAATTAACGATGGTGAAGGA-----TACACAGGGGATATTGTTTTGCT--- 1893
DB 394 LeuThrLeuAsnGlnAlaGluGlyThrLysValTyrAsnGlyArgIleIlePheSerGly 413
QY 1894 -----AATGGAACGAGTACTTTTGTACCAA 1917
DB 414 GluLysLeuThrGluAspGlnThrAlaAspAlaAspAsnLeuLysThrValPheThrGln 433
QY 1918 AATGTTACGATAGCAGGAGGAGGATGTTCTTCTGTAAGGCAAAATTTATCATGAT 1977

```

```

Db      434  ProfileAlaLeuAlaAlaGlyGluLeuLeuLeuAgsnglyValGluValAlaAlaLys 453
QY      1978  TCTTAAGTCACAGAGTGGAGTCTGTATAG---GAAGCTGGAGATACATGCGATTTT 2034
Db      454  AlaValSerGlnThrAlaGlySerLeuLeuLeuLeuLeuAgsAlaGlyThr----- 469
QY      2035  GTAACCTCCACACCACCAAGCCTCTCGCGCTTAATCAGTGTATGATCAGCGCTTCCCAAT 2094
Db      470  -----LysLeuSerAlaLysThrGlnAaspValThrLeuThrLeuThrAen 482
QY      2095  CTGCATTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2154
Db      483  LeuAlaLysProAenSerLeu----- 490
QY      2155  CCTCCAGCGCAAGATCTCATCTCGCAGTCATTTGGTAGC---ACAATGCTGCTGCTGTT 2211
Db      491  -----AaspGlyThrLysIleAlaValIleAlaAlaValAlaAenAlaLysAenVal 507
QY      2212  ACAATTAGTGGCCTATCTTTTGGAGATTTGGATGATACAGCTTATGATAGTATGAT 2271
Db      508  ThrValThrGlyAlaIleGlyIleLeuAaspProThrGlyLysPheThrGluAaspHisLys 527
QY      2272  TGGTAGGTCTTAATCAAAAAATCAATGCTCGAAATTACAGTTAGGACTAAG----- 2325
Db      528  -----LeuAenGluThrLeuAlaLeuGlyGlyIleGlnPheSerAlaLysGlySer 544
QY      2326  ---CCCCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCTAAGTATGCG 2382
Db      545  IleThrThrThrAaspValProSer---ThrThrThrArgSerProAlaGlnHisThrGly 563
QY      2383  TATCAAGGAAGCTGGAGCTTGGCTGG-----GATCCTAATACAGCA 2424
Db      564  TyrGlnGlyAenTrpSerLeuSerTrpIleThrAaspAenGlySerAaspProlysThr 582
QY      2425  AATAATGGTCTTATCTCTGAAAGCTACATGACTGAAACTGGGTATATATCTCGGCTCT 2484
Db      583  -----GlnThrAlaValPheAenTrpAenLysThrGlyThrAenProAenPro 598
QY      2485  GAGCGAGTAGTCTTTTGGTTCCTCAATAGTTATGGGATGCCATTTTAGATATACATCT 2544
Db      599  GluArgArgAlaProLeuValLeuAenSerLeuTrpGlySerPheMetAaspIleArgSer 618
QY      2545  GGCATTACAGCAATTCAGCAAGTGTGGATGGG---CGCTCTTATTTGTCGAGGATATCG 2601
Db      619  IleGlnAaspValMetGluArgSerValAaspThrLeuLeuGluThrArgArgGlyLeuTrp 638
QY      2602  GTTTCCTGGATTCGAAATTTCTCTCATCGCCGCGATGCTTTTAGTCAGCGCATATCGG 2661
Db      639  ValSerGlyValGlyAenPheLeuHisLysAaspProSerAlaGluAenArgLysPheArg 658
QY      2662  TATATTAGTGGGGTATTCTCTTAGGAGCAAACTCCTTACTTTTGA---TCATCATGTTT 2718
Db      659  HisIleSerSerGlyThrValLeuLeuGlyAlaThrThrAenThrSerGlnGluAaspThrLeu 678
QY      2719  GGTCTAGCATTTACCGAAGTATTGGTAGATCTTAAGATTTATGATGCTGCTCTTCCAAT 2778
Db      679  SerValAlaPheCysGlnLeuPheGlyLysAaspLysAaspThrLeuValAlaLysAenAla 698
QY      2779  CATCATGCTTGCATAGGATCCGTTTAT----- 2805
Db      699  AlaAenValThrAlaGlySerIleThrGlnHisValSerLysPheAaspAaspLeuThr 718
QY      2806  -----CTATCTACCCCAACAAAGCTTTTATGTGGATCTCTAT-----TTGTTC 2844
Db      719  ArgLeuPheAenGlyProAenThrCysCysSerGlyPheSerLysGluIleProIlePhe 738
QY      2845  GAGATACGTTTATCCCTGCTAGCTACGGGTTGGGATCAGCATATGAAACCTCATAT 2904
Db      739  LeuAaspAla-----GlnValThrThrCysHisThrAenAenAenMetThrThrThrThr 756
QY      2905  ACATTCAGAGGAGCGATGTTCTGTTGGATATAACTGCTGCGCTGGAGAGATTGGA 2964

```

```

Db      757  ThrAaspTyrProGluValLysGlySerTrpGlyAenAaspThrValGlyValAlaLeuSer 776
QY      2965  GCGGATTTACCGATTTGATGATTACTTCAATGCTTATTTGAATGAGTTCGCTCTTTTC 3024
Db      777  ThrSerValProIleProIlePhe---ThrHisAlaPhePheAaspSerTyrAlaProPhe 795
QY      3025  GTCAAGCTGATGTTTCTTATGTCGCGATCATCAATCTTTTACAGAGAAAGCGATCAAGCT 3084
Db      796  AlaLysLeuGlnValValTyrAlaHisGlnGluAaspPheLysGluProThrArgGluGly 815
QY      3085  CGGCAATTCAGAGCGGACATCTCTTAATCTATCAAGTCTCTGTTGGAGTGAAGTTTGTAT 3144
Db      816  ArgThrPheGluSerSerAaspLeuAenValSerValProIleGlyValLysPheGlu 835
QY      3145  CGATGTTCTAGTACATCTCTAATAATATAGCTTTATGGCGCTTATCTGTGATGCT 3204
Db      836  LysLeuValTyrGlyGlyThrAlaTyrAaspLeuThrLeuMetTyrValProAaspVal 855
QY      3205  TATCGCACCATCTCTGTGTAAGAGACACGCTCTCTATCCCATCAAGACATGGACACA 3264
Db      856  TyrArgHisAenProAenCysIleThrGlyPheAlaIleAenAaspValThrTrpLeuThr 875
QY      3265  GATGCTCTTCTATTAGAACAGCATGGAGTTGTGGTTAGAGGATCTATGTATGCTTCTCTA 3324
Db      876  ThrAlaThrAenLeuAlaArgGlnAlaPheIleIleArgAlaGlyAenHisIleAlaVal 895
QY      3325  ACAAGTATATAGATATATGCGCATCGAAGATATGATATCGAGATGCTTCTCGAGGC 3384
Db      896  ThrSerGlyPheGluMetPheSerGlnPheGlyPheGluLeuArgSerSerArgAen 915
QY      3385  TATGTTTGTGAGTCAGGAAGTAGATCCGGTTC 3417
Db      916  TyrAenValAaspLeuGlyAlaLysValSerPhe 926

RESULT 23
PMP11 CHLPN
ID PMP11 CHLPN STANDARD; PRT; 928 AA.
AC O86164; O9K299;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp11 precursor (Polymorphic membrane
DE protein 11) (Outer membrane protein 4).
GN Name=pmp11; Synonyms=omp4; OrderedLocusNames=CpN0449, CP0302, CpB0468;
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029/VR-1310;
RX MEDLINE=99081766; PubMed=9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer
RL membrane proteins of Chlamydia pneumoniae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RL role in immunopathogenicity.";
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=9906606; PubMed=10192388; DOI=10.1038/7716;
RA Kaiman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RN [4]

```


ID AC Q92393; Q9RB66; STANDARD; PRT; 930 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Probable outer membrane protein pmp8 precursor (Polymorphic membrane protein 8) (Outer membrane protein 11)
 DE Name=pmp8; Synonym=omp11; OrderedLocustNames=Cp00446, CP0307, Cp00463;
 OS Chlamydia pneumoniae (Chlamydia pneumoniae)
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OX NCBI_TaxID=83558;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CWL029 / VR1310;
 RX MEDLINE=20007584; PubMed=10539856;
 RA Christiansen G., Bosen T., Hjerno K., Daugeard L., Mygind P.,
 RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
 RT role in immunopathogenicity";
 RL Am. Heart J. 138:S491-S495(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
 RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Nat. Genet. 21:385-389(1999).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J.D., Uterback T.R., Beery K.J.,
 RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
 RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
 RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.; Chlamydia
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
 RA Shirai M., Hirakawa H., Kinoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=TW-183;
 RA Geng M.M., Schumacher A., Muehlhofer I., Bensch K.W., Schaefer K.P.,
 RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
 RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
 RT other Chlamydia strains based on whole genome sequence analysis";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
 CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AJ133034; CAB37068.1; -; Genomic DNA.
 CC EMBL; AB001627; AAD18590.1; -; Genomic DNA.
 CC EMBL; AB002193; AAF38164.1; -; Genomic DNA.
 CC EMBL; BA000008; BAA98654.1; -; Genomic DNA.
 CC EMBL; AE017158; AAP98394.1; -; Genomic DNA.

DR PIR; A81591; A81591.
 DR PIR; D72078; D72078.
 DR PHCI-2DPAGS; Q92393; -.
 DR TIGR; CP0307; -.
 DR InterPro; IPR005546; Auto.transptbeta.
 DR InterPro; IPR006315; Auto.transpporter.
 DR InterPro; IPR011427; ChlamPMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02415; Chlam_PMP; 6.
 DR Pfam; PF07548; ChlamPMP_M; 1.
 DR TIGRFAMs; TIGR01414; autotransp_barl; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
 KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
 FT SIGNAL 26 Potential.
 FT CHAIN 27 930 Probable outer membrane protein pmp8.
 FT VARIANT 177 177 T -> A (in strain AR39, strain J138 and strain TW-183).
 SQ SEQUENCE 930 AA; 97670 MW; 46A9B5E3BB913C4C CRC64;
 Alignment Scores:
 Pred. NO.: 7.68e-63 Length: 930
 Score: 1079.00 Matches: 303
 Percent Similarity: 43.2% Conservative: 160
 Best Local Similarity: 28.3% Mismatches: 407
 Query Match: 13.7% Indels: 202
 DB: 1 Gaps: 29
 US-10-701-844-1 (1-4435) x PMP8_CHLPN (1-930)
 QY 382 ATGCAAGCTCTTCGATTAAGTCTTCTTCAATGATTCCTAGCTTATCTTGTGCTCT 441
 DB 1 MetLysileProLeuHisLysLeuLeuLeuSerThrLeuVal---ThrProLeu 19
 QY 442 TTAATGGGGGGGATATGACGACGAAATC---ATGTTCTCTCAAGGAATTTACGATGGG 498
 DB 20 LeuSerleuAlaThrTyGlyAlaAspAlaSerLeuSerProThrAspSerPheAspGly 39
 QY 499 GAGAGTTAACTGATTCATTTCCCTATCTGTTTATAGGAGTCGACGATGCTGTT 558
 DB 40 AlaGlyGlySerThrPheThrProLysSerThr---AlaAspAlaAsnGlyThrAsnTyr 58
 QY 559 TTTCTGCGAGGAGTTAACTTAAATAATCTTGCACAAATCTTATGCGATTCGCTTGA 618
 DB 59 ValLeuSerGlyAsnValTyIleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly 77
 QY 619 AGTTGTTTGGAACTTATAGGAGTTTACTGTTTATAGGAGGAGGACACTCCTGCTCACT 678
 DB 78 CysCysPheThrGluThrGlyAspLeuThrPheThrGlyLysGlyTySerPheSer 97
 QY 679 TTCGAGAACATACGACTCTTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGA 738
 DB 98 PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAlaSerThrThrAlaAspLys 117
 QY 739 CTGTTTACTATGAGGTTTAAAGATTTATCTTCCAAATTCGCAATTCATTAATGCTGCC 798
 DB 118 AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe-----Ile 130
 QY 799 GTACTGCTGCTGCAAGCACTAATAAGGTAAGCGAGCTCCGACGACACA----- 849
 DB 131 AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150
 QY 850 TCTACCGCTTAATGCTACTATTATTCTTAAACACAGATCTTTTGTCTACTCAATAATGAG 909
 DB 151 AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal----- 164
 QY 910 AAGTTCTCTATCTATCTTATTTAGTCTCTGGAGATGGGGGAGCTATAGATGCTAAGAGC 969
 DB 165 -----SerAsnGluAlaAsnAsnAsnGlyGlyAlaIleThrThrLysThr 179
 QY 970 TTAACGGTTCAGGAATATAGCAAGCTTTGTGCTTTCAGAGAAATATCTGCTCAAGCTGAT 1029
 DB 180 LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLeu 199

QY	1030	GGGGAGCTTGTCAAGTAGTCACACAGTTTCTCTGCTATGGCTAAACAGAGCTCTCTATTGCC	1089
DB	200	GlyGlyAlaIleTyrSerSerAlaAlaAlaSerIleSerGlyAsnThrGlyGlnLeuVal	219
QY	1090	TTTGATCGGAATGTTGCAGAGTAGAAGGGGGAGGAGTATGCTCTCTCAGATGGCCAG	1149
DB	220	PheMetAsnAsn---LysGlyGluThrGlyGlyGlyAlaLeuGlyPheGlu	235
QY	1150	CAGGGAGTGCATCATCTACTTCAACAGAGATCCAGTAGTAGTAAGTTTTCACAGAAATACT	1209
DB	236	---AlaSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr	252
QY	1210	CGCGTAGAGTTTTCATGGGAACTAGTACCGCAGTAGTAGGAGGAGGATTTACTCTCACGGGAAC	1269
DB	253	AlaThrAsp	255
QY	1270	GTTGCTTTCTGAATAATATGGAAAACCTTGTTCTTCTCAACAATGTTGTTCTCTCTGTTTAC	1329
DB	255	---	255
QY	1330	ATTGCTGCTAAGCAACCAACAAGTGGACAGCTTCTTAATACGAGTAATAATTACGGAGAT	1389
DB	256	AlaAla	259
QY	1390	CGAGGAGCTATCTTCTGTGAAGATGGTGCAGACAGAGATCCAAATCACTCTGATCAGTT	1449
DB	260	GlyGlyAlaIleTyrCysGluLys	274
QY	1450	TCCTTTGATGAGAGGGAGTAGTTTTCTTTAGTAGCAATAGTAGCTGCTGGGAAGGGGA	1509
DB	275	ThrLeSerGlyAsnLysSerLeuThrPheAlaGluAsnSerValThrGlnGlyGly	294
QY	1510	GCTATTATATGCAAAAAGCTCTCGGTGCTCAACTGTGCGCTGTGACAAATTTTAAAGAAAT	1569
DB	295	AlaIleCysAlaHisGlyLeuAspLysSerAlaAlaGlyProThrLeuPheSerAsnAsn	314
QY	1570	ATCGCTAAATGATGTGTGGAGCGATTTATTATTAGAGAGATCTGGAGAG	1614
DB	315	ArgCysGlyAsnThrAlaAlaGlyGlyAlaIleAlaIleAlaAspSerGlySer	334
QY	1615	CTCAGTTTATCTGCTGATTATGGAGATATATTTTCGATGGGAATCTTAAATGAT	1674
DB	335	LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn	349
QY	1675	AAAGAGAAATGTGCCGATGTTAATGGCGTAACCTGTGCTCTCAAGCCATTTCGATGGGA	1734
DB	350	ThrLeuThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyrLeuGly	367
QY	1735	TCGGAGGGAAATTAACACATTAAGACTAAAGCAGGGCATCAGATCTCTTTAATGAT	1794
DB	368	SerSerAlaLysIleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyrPheTyrAsp	387
QY	1795	CCCATCGAGATGGCAACCGAAAT	1833
DB	388	ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer	407
QY	1834	TCTTCCAACTTCTAAAAATTAAACGATGCTGAAGGATACACAGGGGATATGTTTGTCT	1893
DB	408	AsnSerProLeu	420
QY	1894	ATGGAAGCAGTACTCTGTAC	1914
DB	421	GlyGluLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys	440
QY	1915	CAAAATGTTTACGATCAGCAGCAAGGAGTGTCTTCTTCCGTAAGGAGCAAAATTCATG	1974
DB	441	GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLeuLysGlyAsnValGluLeuAspVal	460
QY	1975	AATTCTTAAGTCACAGGTGGAGTCTGTATATATGGAAGCTGGGAGTACATGGGATTTT	2034
DB	461	AsnGlyPheThrGlnThrGluGlySerThrLeuLeu	472

QY	2035	GTAACTCCCAACACCACCAACAGACCTCTCTGCGGTAAATCACTGATCATCGCTTTCCTCAAT	2099
DB	473	-----MetGlnProGlyThrLysLeuLysAlaAAspThrGluAlaLeuSerLeuLys	490
QY	2095	CTGCATTGTGCTCTTCTCTTTGTTGAGCAAAACAATGCGAGTTACGAAATCCTCTACCAAT	2154
DB	491	LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer	505
QY	2155	CCTCCAGCGCAAGATTCTCATCTCTGCGCATTTGGTAGCACAACTGCTGCT	2205
DB	506	-----IleGluThrAlaGlyAlaAsnLys	513
QY	2206	TCGTGTACAATTAGTGGGCTATCTTTTGTGAGGATTTGGATGATACAGCTTATGATAG	2265
DB	514	ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrGlu	532
QY	2266	TATGATTGGCTAGTCTTCAATCAAAAANAATCAATGTCTCTGAAATTACAGTTAGGAGCAATAAG	2325
DB	533	-----SerHisThrIleAsn	542
QY	2326	CCC-----CCAGCTAATAGCCCCATCAGATTTGACTCTAGGG	2361
DB	543	ProLeuValValPheThrAlaAlaThrAlaAlaSerAspIleTyrIleAspAlaLeuLeu	562
QY	2362	-----AATGAGATGCCTAAGTATGCTTATCAAGAGAGCTGGAAGCTTGGC	2406
DB	563	ThrSerProValGlnThrProGluProHisIleTyrGlyTyrGlnGlyHisIleTyrGluAlaThr	582
QY	2407	TGG-----GATCTCAATACAGCAAAATAATGGTCTTATATCTCTGAAAGCTACATCGACTAAA	2463
DB	583	TrpAlaAspThrSerThrAlaLysSerGly	598
QY	2464	ACTGGGTATATCTCTGGGCTGAGCGAGTAGTCTTCTTGGTTCCAAATAGTTATGGGGA	2523
DB	599	ThrGlyTyrAsnProAsnProGluArgArgAlaSerValValProAspSerLeuTrpAla	618
QY	2524	TCCATTTTAGATATACAGATCTGGCATTCAGCAATCAACGCAAGTGTGGATGGCGGCTCT	2583
DB	619	SerPheThrAspIleLeuThrLeuGlnGlnIleMetThrSerGlnAlaAsnSerIleTyr	638
QY	2584	TATTGTCGAGGATTTGGTTCTCGGAGTTTCGAATTTCTTCTATCATGACCCGCGATGCT	2643
DB	639	GlnGlnArgGlyLeuTrpAlaSerGlyThrAlaAsnPhePheHisLysAspLysSerGly	658
QY	2644	TTAGTCTCAGGATATCGGTATATTAGTGGGGTTATTCTTCTTAGGAGCAAACTCTCTAC	2700
DB	659	ThrAsnGlnAlaPheArgHisLysSerTyrGlyTyrIleValGlyGlySerAlaGluAsp	678
QY	2701	TTTGATCATCGATTTGGTCTAGCATTTTACCGAGTATTTGGTAGATCTTAAGATTTAT	2760
DB	679	PheSerGluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysAspLeu	698
QY	2761	GTAGTGTCTGTTCCAAATCATCGCTTGATAGGATCCGTTTATCTATCTACCCACAA	2820
DB	699	PheIleValGluAsnThrSerHisAsnTyrLeuAlaSerLeuTyrLeuGlnHisArgAla	718
QY	2821	GCTTTATGTGATCTCATTTTG-----TTCCGAGATCGGTTT	2856
DB	719	PheLeuGlyGlyLeuProMetProSerPheGlySerIleThrAspMetLeuLysAspIle	738
QY	2857	-----ATCCGTGCTAGTACGGGTTTGGGAATCAGCATATGMAAACCTCA	2901
DB	739	ProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThrArg	758
QY	2902	TATACATTGTCAGAGGAGCGATGTTGTTGGGATATTAATCTCTCTGCTGGAGAGATT	2961
DB	759	TyrThrSerTyrProGluAlaGlnGlySerThrAsnAsnSerGlyAlaLeuGluLeu	778
QY	2962	GGAGCGGATTAACCATGTCATTAATCTCCATTAAGCTCTATTTGAATAGTTCGCTCT	3021
DB	779	GlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPhePheGlnGlyTyrPhePro	798
QY	3022	TTTCGTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACGAGGAAGCGCATCA	3081

```
Db 799 PheLeuIysPheGlnAlaValTyrSerArgGlnGlnAsnPheLysGluSerGlyAlaGlu 818
Qy 3082 GCTCGGCAATCAAGAGCGGACATCTCTAAATCTATCATGTTCTCTGGAGTGAAGTTT 3141
Db 819 AlaArgAlaPheAspGlyAspLeuValAenCysSerIleProValGlyIleArgLeu 838
Qy 3142 GATCGATGTTCTAGTACATCCTCAATATAATATAGCTTTATGGCGCTTATATCTGTGAT 3201
Db 839 GluIysIleSerGluAspGluLysAsnAenPheGluIleSerLeuAlaTyrIleGlyAsp 858
Qy 3202 GCTTATCGCACCATCTCTGGTGTGAGACACAGCTCTCTATCCATCAAGACATGGACA 3261
Db 859 ValTyrArgLysAenProArgSerArgThrSerLeuMetValSerGlyAlaSerIlePhe 878
Qy 3262 ACAGATCGCTTTCAATTAGCAAGACATGAGTGTGTGTAGAGGATCTATGATGCTTCT 3321
Db 879 SerLeuCysLysAenLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeuThr 898
Qy 3322 CTACACAGTATATAGATATATGGCCATCGGAGATATGATATGATCGATGCTTCTCGA 3381
Db 899 LeuSerProHisValGluLeuSerGlyGluAlaTyrGluLeuArgGlySerAlaHis 918
Qy 3382 GCCTATGTTTGTAGTGCAGGAGTAGAGTCCGGTTC 3417
Db 919 IleTyrAsnValAspCysGlyLeuArgTyrSerPhe 930

RESULT 25
Q5L6J2_CHLAB PRELIMINARY; PRT; 1378 AA.
AC Q5L6J2
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polymorphic outer membrane protein.
GN Name=omp15G; Synonyms=pomp145A; OrderedLocusNames=CAB283;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83555;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S26/3;
EX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingston M., Cordero-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrall B.G., Parkhill J., Longbottom D.,
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation."
RL Genome Res. 15:629-640(2005).
DR EMBL; CR48038; CAH63733.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 1378 AA; 144965 MW; 40B2108E7F26BP89 CRC64;

Alignment Scores:
Pred. No.: 1.16e-62 Length: 1378
Score: 1077.00 Matches: 330
Percent Similarity: 41.1% Conservative: 178
Best Local Similarity: 26.7% Mismatches: 382
Query Match: 13.7% Indels: 346
DB: 2 Gaps: 42

US-10-701-844-1 (1-4435) x Q5L6J2_CHLAB (1-1378)
Qy 499 GAGAGCTTAAGTGTATCTTCCCTATATCTGTATA----- 534
Db 226 GluThrLeuGluSerSerLeuProHisThrProLeuSerPheThrAlaSerSerAspVal 245
Qy 535 -----GGAGATCCGAGTGGGACTACTGTTTTCGTCAGGAGAGTTAACATTA 582
Db 246 AlaAenGluHisGlySerProSerProGluThr----- 256
Qy 593 AAAAATCTTGACAAATCTATTGCGAGT-----TTGCCCTTTAAGTTGTTTT 627
```

```
Db 257 LysThrProThrAlaValAlaSerGlnAlaAsnGlyGluThrAsnLysSerCysPhe 276
Qy 628 GGGAACTATTAGGGAGTCTTTACTGTTTATGGGAGAGGACACACTGTTGACTTTTCGAGAAC 687
Db 277 SerAsnThrAspGlySerLeuThrPheValGlyIleAsnHisSerLeuThrPheSerAen 296
Qy 688 ATACGGACTTCTACAAATGGGAGCTCTAATGCTAATAGCCCTGCTGATGAGTCTTTACT 747
Db 297 IleSerValThrAlaProGlySerAlaIleAsnAsnSerAlaGlySerAlaLeu---Thr 315
Qy 748 ATTGAGGCTTTTAAAGAAATATCTCTT-----TCCAATTGC 783
Db 316 PheSerGlyPheLysAspLeuSerPheValCysAlaThrAsnGlnAspGlnThrLysAla 335
Qy 784 AATTCAATTACTGCGTACTGCTGCTGCAACGACTAATAAGGCTAGCCAGACTCCGACG 843
Db 336 AspSerAlaIleTyrValGlyProLysAlaThrSerGlyAspGlyGluLeuSerValGln 355
Qy 844 ACAACATCTACAGCTCTAATGCTACTATTATTTCTAAACAGATCTTTGTTTACTCAAT 903
Db 356 AlaGluGlySerLysAlaAsnSerHisThrGluGlySerValAsnThrAlaLeuIleLys 375
Qy 904 AATGAGAAGTCTCATCTTATAGTAAATTTAGTCTCTCGAGATGGGAGCTATAGATGCT 963
Db 376 -----TyrAlaGly---IleProAlaAspAlaThrAlaThrAsnGly 388
Qy 964 AAGAGCTTAAACGGTTCAAGGAATTTAGCAAGCTTTGTGTCTTCCAAGAA----- 1011
Db 389 GluSer-----GlnGluAspProAlaThr 396
Qy 1012 AATTACTGCTCAAGCTGATGGGAGGCTGTCAAGTACTACAGTCTTCTCTGCT----- 1065
Db 397 AsnValAlaProAlaThrGlnGlyGlyAlaGluLysGluGluAlaAlaAsnSerProAsp 416
Qy 1066 -----ATGGCTAACGAGGCTCTATTGCTCTTGTAGCGAATGTTGCGAGGAGTAAAGAGG 1119
Db 417 IleIleLeuLysGlnAsnValAsnIleThrPheSerSerAsnSerSerLysThrAlaGly 436
Qy 1120 GGAGGAGTGTGCTGTGTAGGATGGGAGGAGTGTGCATCATCTACTTCAACAGAA 1179
Db 437 GlyAlaIleArgVal-----SerGlySerAlaThrIleGluAsnAsnThrGlyThr--- 453
Qy 1180 GATCCAGTATAGTATTTTCCAGAAATACTCGGTGAGAGTTGATGGAAAGTATGCCGA 1239
Db 454 -----CysThrPheSerAsnAsnAlaLysGlu----- 463
Qy 1240 GTAGGAGGAGGATTTACTCTACGGGAAAGTGTCTTCTCTGAATAATGAAAAACCTTG 1299
Db 464 GlnGlyGlyAlaIleSerValAsnGlyAsnCysAspIleThrGlyAsnLysAsnValVal 483
Qy 1300 TTTCTCAACAATGTTGCTCTCTCTCTGTT-----TACATT 1332
Db 484 PheSerGlyAenGlnAlaGlnGluIleProAlaProSerIleValThrValGluAla 503
Qy 1333 GCTCTAAGCAACCAACAGTGGACAGGCTCTTAATAGGAGTAAATAATTACGAGATGGA 1392
Db 504 AlaValGluGluProValGluAlaLys-----GlySerGly 515
Qy 1393 GGAGCTATCTCTGT----- 1407
Db 516 GlyAlaIleTyrCysValLysAlaProIleSerIleProAlaThrProIleGlnProPro 535
Qy 1408 -----AAGAATGGTCGCAAGCAGGATCCAAATACTCT----- 1440
Db 536 LysGlnThrLeuProSerLeuLysSerAlaAlaLeuSerGlyGluThrHisAlaLysVal 555
Qy 1441 -----GGATCAGTTTCTTGTGATGAGGAGGAGTGTAGTT 1473
Db 556 AlaGlnLysGluGluProSerProAspProCysLeuThrIleSerGlyAsnAlaSerVal 575
Qy 1474 TTCTTTAGTACCAATGATGCTGCTGGGAAAGGGAGCTATTATATGCAAAAAGCTCTCG 1533
```

576 IlePheAspAsnAsnSerSerThrValThrGlyGlyAlaIleHisAlaLysLysValVal 595
1534 GTTGCTAACTGGCCCTGTACAAATTTTAAAGGAATATCGTAATGATGGAGCGATT 1593
596 LeuSerSerGlyAsnMetThrPheSerAsnAsnSerSerGlyLysGlyAlaIle 615
1594 TATTTAGGAGATCTGGAGCTCAGTTTATCTGCTGATTAAGAGATATATTTTCGAT 1653
616 TyrIleAlaAspGlyGlyAspIleSerIleThrAlaThrGlySerIleIlePheGln 635
1654 GGGAACTTTAAAGAACAGCCAAAGAGAT 1683
636 GlyAsn---LysValThrAlaAlaAspSerIleThrLeuProThrLysLysGluIleAla 654
1684 -----GCTGCCGATGTTAATGGCTAAGTCTGCTCACAAGCC----- 1722
655 IleAlaAlaGluSerIleGluGlyThrAspProSerGlnAlaSerGlySerAlaSer 674
1722 ----- 1722
675 AlaGlyMetThrSerAlaPheThrLeuAlaAsnLysAlaGluIleProAlaGluSer 694
1723 -----ATTTCGATGGATCGGAGGGAAA 1746
695 GlnAlaLysGluAsnSerLysProThrCysAsnSerIleHisLeuGlySerGlyAlaLys 714
1747 ATAACGACATTAAGAGCTAAAGCAGGCGATCAGATCTCTTAAATGATCCCATC----- 1800
715 IleSerGlnLeuArgAlaGlnThrGlyGlnThrIlePhePheTyrAspProIleThrThr 734
1801 -----GAGATGGCAAC 1812
735 ThrAlaProAlaAlaAlaValValThrAlaLysGlnProGluAlaSerLeuAlaLys 754
1813 CGAATAAAC-----CAGCCAGCGCAGTCT 1836
755 AlaThrSerGlyIleProAlaSerAlaValSerValProAlaProAlaProThr 774
1837 TCCAACTCTTAAATTAAC----- 1857
775 LysThrProLeuLysIleAsnAlaProAspThrGlnAspProGluIleGlnLysValAla 794
1858 -----GATGGTGAAGGATACACAGGGATATCTTTTGTCT----- 1893
795 AlaGluAlaAlaGlnGlnSerAlaValTyrAsnGlyLysIleValPheSerGlyGluLys 814
1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTT 1923
815 LeuSerSerGluAspAlaLysAsnProLeuAsnAlaThrSerValIleHisAsnAspVal 834
1924 ACGATAGACAAAGAGATGTTCTTCGTGGAAGGCAAAATATCAGTGAATCTCTA 1983
835 SerLeuGluAlaGlyThrLeuValLeuSerSerGlyAlaGlyLeuLeuValAspSerPhe 854
1984 AGTCAGACAGCTGGAGTCTGTATATGGAAGCTGGAGTACATGGGATTTT-----GTA 2037
855 ThrGlnLysGluGlySerLeuIleValMetAspGlyGlyThrSerIleIleThrAsnVal 874
2038 ACTCCA-----CAACCACCAACAGCCTCCTGCC----- 2067
875 ThrProAlaSerGluGlyLeuGlnSerArgSerThrProProSerProLysAsnAlaIle 894
2068 -----GCTAATCAGTTG----- 2079
895 ProValIleArgAlaValSerLysValIleAlaSerSerLeuIleAsnLeuArgGluArg 914
2080 -----ATC 2082
915 AlaAspSerGlyAlaGlyAlaValProThrIleGluGluSerProAspGlySerIle 934
2083 ACGTTTCCAACTGCAATTTGCTCTTTCTTTGTTAGCAACATGAGTACGAT 2142
935 ThrIleThrAsnLeuAlaValAsnLeuAspSerLeuGluAsnGlyLysValIleThr--- 953

2143 CCTCTACCAATCTCCAGCGCAAGATTCTCATCTCGCAGTCATTGTGTAGCACAACTGCT 2202
954 -----LeuAlaAlaLysGlySer 960
2203 GGTCTCTTCAATAGTGGCCCTATCTTTTGGAGGATTTGGATGATACAGCTTATGAT 2262
961 GlySerValThrLeuThrGlyAspGlnPheGlnAspSerSerGlnAsnPheTyrAsp 980
2263 AGGTATGATGGCTAGGTTCTTAATCAAAAATCAATGCTCGMAATTAACAGTTAGGACT 2322
981 Asn---ProLeuLeuAsnLysAsnPheThrLeuAsnPheLeuAspIle-----Ser 996
2323 AAGCCCCAGCTAAT-----GCCCATCAGATTGACT 2355
997 AlaProAspAlaGluLysIleHisThrGluGlyPheAsnIleIleProGlnGlyAlaThr 1016
2356 CTAGGAATGATAGCTTAAGTATGGCTATCAGGAAGCTGGAAGCTTGGCTGGATCCT 2415
1017 SerSerAsn-----LeuGlyTyrGlnGlyLysTrpGluValThrGluValLys 1032
2416 AATACAGCAATATGCTTCTTATCTCTGAAGCTACATGAGCTAAAACTGGGTATAAT 2475
1033 AspSerSerGlyLysValSerPheGluMetLys-----TrpValSerThrGlyTyrIle 1050
2476 CTGGGCTGAGCGAGTAGTCTTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGAT 2535
1051 ProThrAlaAsnArgAlaThrLeuValProAsnSerValTyrCysSerAlaIleAsp 1070
2536 ATACATCTGGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGTCGAGA 2595
1071 MetArgAlaPheGlnAsnLeuValGluValSerThrGluGlyGluAspPheHisArgIly 1090
2596 TTATCGGTTTCTGGAGTTTCGAATTTCTTATCATGACCGCGATCTTTAGTCAGGGA 2655
1091 LeuTrpIleSerGlyIleSerAsnPheHisLysAspSerThrLysValGlnGluGly 1110
2656 TATCGGTATATTAGTGGGGTTATCTCTTAGGACAAACTCCTAC-----TTTGGATCATCG 2712
1111 PheArgHisIleSerSerGlyTyrValValGlyValSerThrGluProIleSerAsnLys 1130
2713 ATGTTTGGTCTAGCATTTACGAAGTATTTGTAGATCTTAAAGATATATAGTGTGCTGT 2772
1131 ValMetAspLeuAlaPheCysGlnMetLeuGlyLysSerLysAspTyrArgLeuAlaAsp 1150
2773 TCCATCATCATCTGTCATAGATCCGTTTATCTATCTACCAACAAGCTTTATGTGGA 2832
1151 AlaArgSerHisValTyrAlaAlaSerIleHisThrLysCysGluLys---LeuValAsn 1169
2833 TCCTATTCTTCTCGGA----- 2847
1170 HisThrPheSerLysArgLysGlyAlaIleLeuAlaArgLysProGluLysSerPro 1189
2848 -----GATCGGTTTATCCGTCTAGCTACGGTTGGGAATCAGCATATGAAAACC 2898
1190 IleIlePheAspAla-----GlnLeuSerTyrSerLeuSerHisAsnSerMetThrThr 1207
2899 TCATATACATTTGTCAGAGAGCGATGTTTGTGGGATAATACTGTCTGCTGCGAGAG 2958
1208 LysHisThrProAsnProSerArgGlyLysTrpAsnAsnHisCysValAlaGlyGlu 1227
2959 ATTGGAGCGGATTACCGATTGATTACT---CCATCTAAGCTCTATTGTAATGAGTTG 3015
1228 LeuGlySerTyrLeuProIleLeuValAspHisProIle-----IleGluGluLeu 1244
3016 CGTCTCTTCTCGTCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTT-----ACAGAG 3069
1245 PheProPheValLysLeuHisIleValPheValGlnGlnGluAspPheLysGluThrGln 1264
3070 GAAGCGCATCAAGCTCGGCAATTCAGAGCGGACATCTCTTAATCTATCAGTCTCTGTT 3129
1265 GlyGlyThrGluAsnArgAsnPheGlnSerAlaHisPheValAsnValSerLeuProLeu 1284


```

Db      690 GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLyAsnHisGlyAspThrTyGly 709
Qy      2794 GGATCCGTTTATCTATCTACCAACAACAGCTTTA-----TGTGATCTCTATTGTTGGGA 2847
Db      710 AlaSerLeuTyPheHisThrGluGlyLeuPheAlaAsnHisLeuTrpGly 729
Qy      2848 GATCGTTT----- 2856
Db      730 LysAlaThrArgAlaProTrpValLeuSerGluLeuSerGlnIleIleProLeuSerPhe 749
Qy      2857 ---ATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTGCA 2913
Db      750 AspaLalysPheSerTyLeuHisThrAspAanHisMetTyThrTyThr----- 767
Qy      2914 GAGGAGCGGATGTCGT-----TGGATTAATAACTCTCTGGCTGGGAGAGATTGGACG 2967
Db      768 AspAanSerIleIleLysGlySerTrpArgAanAspaLalysPheCysAlaAspLeuGlyAla 787
Qy      2968 GGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGAGTTGGTCCTTTCGTG 3027
Db      788 SerLeuProPheValIleSerValProTyLeu---LeuLysGluValGluProPheVal 806
Qy      3028 CAAGCTGAGTTTCTTATGCGCATCATCAATCTTTTACAGAGGAGGCGATCAAGCTCGG 3087
Db      807 LysValGlnTyIleTyAlaHisGlnGlnAspPheTyGluArgTyAlaGluGlyArg 826
Qy      3088 GCATTCAAGAGCGGACATCTCTAAATCATCATGTTCTCTGTGGAGTGAAGTTTGATGCA 3147
Db      827 AlaPheAenLysSerGluLeuIleAenValGluIleProIleGlyValThrPheGluArg 846
Qy      3148 TGTTCTAGTACATCCTCAATAAATAGCTTTATGGCGGCTTATCTGTCGATGCTTAT 3207
Db      847 AspSerLysSerGluLysGlyThrTyAspLeuThrLeuMetTyIleLeuAspAlaTy 866
Qy      3208 CGCCACCATCTCTGTACTGAGACACAGCTCTATCCATCAAGACATGACACACAGAT 3267
Db      867 ArgAanProLysCysGlnThrSerLeuIleAlaSerAspAlaAenTrpMetAlaTy 886
Qy      3268 GCCTTCATTAGCAAGCATGAGTGTGTGTAGAGATCTATGATGCTTCTCTAACA 3327
Db      887 GlyThrAenLeuAlaArgGlnGlyPheSerValArgAlaAlaAenHisPheGlnValAen 906
Qy      3328 AGTAATATAGATATATGCGCATCGGATGATGATGATGATGATGATGATGATGATGAT 3387
Db      907 ProHisMetGluIlePheGlyGlnPheAlaPheGluValArgSerSerArgAanTy 926
Qy      3388 GGTGTGAGTCAGGAAGTAGATCGGTTTC 3417
Db      927 AsnThrAenLeuGlySerLysPheCysPhe 936

RESULT 27
PMP9 CHLPN
ID PMP9 CHLPN STANDARD; PRT; 928 AA.
AC Q92398;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmp9 precursor (Polymorphic membrane protein 9) (Outer membrane protein 10).
GN Name=pmp9; Synonyms=omp10; OrderedLocusNames=CP0447, CP0306, CP0464;
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=2007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity."
RL Am. Heart J. 138:8491-8495(1999).
RN [2]

```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.P.,
RA White O., Hickey E.K., Peterson J.D., Ueberback T.R., Barry K.J.,
RA Bass S., Linher K.D., Weidman J.P., Khouri H.W., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Bisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RX Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -1- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ133034; CAB37069.1; -; Genomic DNA.
CC EMBL; AE001628; AD18591.1; -; Genomic DNA.
CC EMBL; AE002392; AAF38163.1; -; Genomic DNA.
CC EMBL; BA000008; BAA98655.1; -; Genomic DNA.
CC EMBL; AE017158; AAF98395.1; -; Genomic DNA.
CC PIR; B72077; B72077.
CC PIR; B86546; B86546.
CC TIGR; CP0306; -.
CC InterPro; IPR005546; Auto_transptbeta.
CC InterPro; IPR006315; Autotransporter.
CC InterPro; IPR011427; ChlamPMP_M.
CC InterPro; IPR003368; Chlamydia_PMP.
CC InterPro; IPR006626; Pbh1.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; Chlam_PMP; 7.
CC Pfam; PF07548; ChlamPMP_M; 1.
CC SMART; SM00710; Pbh1; 5.
CC TIGRFAMs; TIGR01414; autotrans_bar1; 1.
CC TIGRFAMs; TIGR01376; POMP_repeat; 6.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 928 Probable outer membrane protein pmp9.
SQ SEQUENCE 928 AA; 98333 MW; 58910A8F04F12219 CRC64;
Alignment Scores: 5.46e-59 Length: 928
Pred. No.:

```

Score: 1021.00 Matches: 313
Percent Similarity: 44.6% Conservative: 161
Best Local Similarity: 29.5% Mismatches: 404
Query Match: 13.0% Indels: 184
DB: 1 Gaps: 39

US-10-701-844-1 (1-4435) x PMP9_CHLPN (1-928)

QY 382 ATGCAAACTCTTCCATAGATCTCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441
DB 1 MetLysSerLeuHisTrpPheLeuIleSerSerLeuAlaLeuPro---LeuSer 19
QY 442 TTAATATGGGGGGATATCAGCAGAAATCATGGTT-----CCTCAAGGAATTTAC 492
DB 20 LeuAsnPheSerAlaPheAlaValValGluIleAsnLeuGlyProThrAsnSerPhe 39
QY 493 GATGGGAG---ACGTTAACTGTATCATTTCCCTATATCTGTATAGAGATCCGATGGG 549
DB 40 SerGlyProGlyThrTyrThr-----ProProAlaGlnThrThrAsnAlaAspGly 56
QY 550 ACTACTGTTTTCTGCAGAGAGTAAATTAATAAAATCTTGACAAATTTCTATTCAGCT 609
DB 57 ThrIleTyrAsnLeuThrGlyAspValSerIleThrAsnAla---GlySerProThrAla 75
QY 610 TTGCTTTAAGTTGTTTGGAACTTATTAGGAGTTTACTGTTTTCAGGAGAGACAC 669
DB 76 LeuThrAlaSerCysPheLysGluThrThrGlyAsnLeuSerPheGlnGlyHisGlyTyr 95
QY 670 TCGTTGACTTTCAGAAACATACCGACTTCTACAAATGGGCGAGCTCTAAGTAATACGCT 729
DB 96 GlnPheLeuLeuGlnAsnIleAspAlaGlyAlaAsn---CysThrPheThrAsnThrAla 114
QY 730 GCTGATGGAAGTTTACTATGAGGGTTTAAAGAAATATCTTTTCCAAATGCAATTCA 789
DB 115 AlaAsnLysLeuSerPheSerGlyPheSerTyrLeuSerLeu----- 129
QY 790 TTAATTGCCGTACTGCTGTCACAGCATTAATAAGGGTAGCCAGACTCCGACGACAA 849
DB 130 -----lleGlnThrThrAsn 134
QY 850 TCTACACCGTCTAATGCTACTATTTATTTAAACACAGATCTTTTGTACTCAATAATAG 909
DB 135 AlaThrThrGlyThrGlyAlaIleLysSerThrGlyAlaCysSerIleGlnSerAsnTyr 154
QY 910 AAGTTCTCATTTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAC 969
DB 155 SerCysTyrPheGlyGlnAsnPheSerAsnAspAsnGlyGlyAlaLeuGlnGlySer 174
QY 970 TTAACGTTCAAGGATACCAAGCTTTGCTCTCCAAAGAAATCTGCTCAAGCTGAT 1029
DB 175 IleSerLeu---SerLeuAsnProAsnLeuThrPheAlaLysAsnLysAlaThrGlnLys 193
QY 1030 GGGGAGCTGTGCAAGTAGTCACAGTGTCTCTGCTATGCTAAGCTAACAGGCTCTATTGCC 1089
DB 194 GlyGlyAlaLeu-----TyrSerThr----- 200
QY 1090 TTTGTACGGAATGTTGACGAGTAGAAGGGGGAGGATGCTGCTGTTCCAGATGGGCGAG 1149
DB 201 -----GlyGlyIleThr----- 204
QY 1150 CAGGAGTGTCAATCTACTTCAACAGAGATCCAGTAGTAAAGTTTTCAGAAATACT 1209
DB 205 -----IleAsnAsnThrLeuAsnSer-----AlaSerPheSerGluAsnThr 218
QY 1210 GCGGTAGAGTTTGAGGAGCTAGTCCGAGTAGGAGGAGGATTTACTCTACGCGGAAC 1269
DB 219 Ala-----AlaAsnAsnGlyGlyAlaIleTyrThrGluAlaSer 231
QY 1270 GTTGCTTTCTGGAATAATGGAACCTTG---TTTCTCAACATGTTGCTTCTCTGTT 1326
DB 232 ---SerPheIleSerSerAsnLysAlaIleSerPheIleAsnAsn----- 245
QY 1327 TACATTGCTGTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATTACGGA 1386

DB 246 SerValThrAlaThrSerAlaThr----- 253
QY 1387 GATGGAGAGCTATCTTCTGTAAAGATCGTGGCAAGCAGGATCCAATAACTCTGGATCA 1446
DB 254 ---GlyGlyAlaIleTyrCys-----SerSerThrSerAlaProLysProVal 268
QY 1447 GTTTCCTTTGATGGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGGGGAAGGG 1506
DB 269 LeuThrLeuSerAspAsnGlyGluLeuAsnPheIleGlyAsnThrAlaIleThrSerGly 288
QY 1507 GGAGCTATTTATGCCAAAAGCTCTCGGTTGCTAACTGTGCCCCGTGTACAAATTTTAAGG 1566
DB 289 GlyAlaIleTyrThrAspAsnLeuValLeuSerSerGlyProThrLeuPheLysAsn 308
QY 1567 AATATCGCTTAATGAT-----GGTGAGCGATTATTATTAGGAGAATCTGGA 1611
DB 309 AsnSerAlaIleAspThrAlaAlaProLeuGlyGlyAlaIleAlaIleAlaAspSerGly 328
QY 1612 GAGCTCAGTTATCTGCTGATTATGAGATATTTATTTTCATGGGAATCTTAAAGAAC 1671
DB 329 SerLeuSerLeuSerAlaLeuGlyAspIleThrPheGluGlyAsn-----Thr 345
QY 1672 GCCAAGAGAAATGCTCCGATGTTAATGGCGTAACGTGTCTCTCACAGCCATTTCCATG 1731
DB 346 ValValLysGlyAlaSer-----SerSerGlnThrThrThrArgAsnSerIleAsnIle 363
QY 1732 GGA---TCGGGAGGAGAAATACGACATTAAGAGCTTAAAGCAGGCGATCAGATTCTCTT 1788
DB 364 GlyAsnThrAsnAlaLysIleValGlnLeuArgAlaSerGlnGlyAsnThrIleTyrPhe 383
QY 1789 AATGATCCCATCGATGGCAACGGAATAACACGACCCAGCGCATCTTCCAAACTCTTA 1848
DB 384 TyrAspProIleThr-----ThrSerIleThrAlaAlaLeuSerAspAlaLeu 399
QY 1849 AAAATTAACGATGGTGAA-----GGATACACAGGGAGTATGTTTGTGCT 1893
DB 400 AsnLeuAsnGlyProAspLeuAlaGlyAsnProAlaTyrGlnGlyThrIleValPheSer 419
QY 1894 -----AATGGAAGCAGTACTTTGTAC 1914
DB 420 GlyGluLysLeuSerGluAlaGluAlaGluAlaAspAsnLeuLysSerThrIleGln 439
QY 1915 CAAAATGTTACGATAGCAAGGAGATTTCTTCGTCAAAAGGCAAAATATATCAGTG 1974
DB 440 GlnProLeuThrLeuAlaGlyGlnLeuSerLeuLysSerGlyValThrLeuValAla 459
QY 1975 AATTTCTTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGAT 2031
DB 460 LysSerPheSerGlnSerProGlySerThrLeuLeuMetAspAlaGlyThrThrLeuGlu 479
QY 2032 TTTGTAACCTCCACACCACCAACGACCTCTCCGCGCTAATCAGTTGATCAGCTTTCC 2091
DB 480 ThrAlaAspGly-----lleThrIleAsn 487
QY 2092 AATCTCATTTGCTCTTCTTCTTTGTTAGCAACAATGCAGTTACGAATCCTCTTACC 2151
DB 488 AsnLeuValLeuAsnValAspSerLeu----- 496
QY 2152 AATCTCCAGCGCAAGATTTCTCATCTCGCAGTCATTTGGTAGCAACAATGCTGCT 2208
DB 497 -----LysGluThrLysLysAlaThrLeuLysAlaThrGlnAlaSerGlnThr 512
QY 2209 GTTCAATTAGTGGGCTATCTTTTTTAGAGATTTCAGTATGATGATAGGTAT 2268
DB 513 ValThrLeuSerGlySerLeuSerLeuValAspProSerGlyAsnValTyrGluAspVal 532
QY 2269 GATTGGCTAGTTCTTAATCAAAAATCAATGTCTCTGAAATTTACAGTTAGGAGTAAGCCC 2328
DB 533 SerTrp-----AsnAsnProGlnVal---PheSerCysLeuThrLeuThrAlaAspAsp 549
QY 2329 CCAGCTAATGCCCA---TCAGATTGTGACT---TAGGGAATGATGCTTAAAG 2376


```
DR PHCI-2DPAGE; O9Z899; ..
DR TIGR; CP0309;
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR006626; Pbh1.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 14.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR SMART; SM00710; Pbh1; 10.
DR TIGRPFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRPFAMs; TIGR01376; POMP repeat; 13.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 23 Potential.
FT CHAIN 24 1276 Probable outer membrane protein pmp6.
FT CONFLICT 421 421 E -> EGCGAILAFIDSGVSDKLTGLSIANNOEVLSTNSA
FT ATVGSGIYVTFTSTNTAKTGGALYSKGNNSLSGNTNLLPSG
FT PTLTSGTSPVTSNANOE (in Ref. 1 and 4).
FT NKATGPSNNSANOE (in Ref. 1 and 4).
SQ SEQUENCE 1276 AA; 132127 MW; C97FA40662C88E0B CRC64;

Alignment Scores:
Pred. No.: 1,49e-58 Length: 1276
Score: 1015.00 Matches: 342
Percent Similarity: 37.9% Conservative: 152
Best Local Similarity: 26.3% Mismatches: 421
Query Match: 12.9% Indels: 38
DB: Gaps: 38

US-10-701-844-1 (1-4435) x PMP6_CHLPN (1-1276)
QY 493 GATGGGAGAGCTTAAGTATCAATTCCTATATAGGAGATCGAGTGGACT 552
Db AsnGlySerSerGlySerAlaAlaPheThrAlaLysGluThrSerAspAlaSerGlyThr 55
QY 553 ACTGTTTTTTCGACGAGAGTTAAACATTAATAAATCTTGACAAATCTATTGCGCTTTG 612
Db ThrTyThrLeuThrSerAspValSerIleThrAsnVal---SerAlaIleThrProAla 74
QY 613 CCTTTAAGTTGTTTGGGAACCTATTAGGAGTCTTACTGTTTATAGGAGAGACACTCG 672
Db AppLysSerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSer 94
QY 673 TTGACTTTTCGAGAACATACGACTTCTCAAAATGGGGAGCTCTAAGTATATACCGCTGCT 732
Db LeuValLeuGlnThrIleAlaLeuThrHisAspGlyAlaAlaIleAsnAsnThrAsnThr 114
QY 733 GATGGACTGTTTACTATTGAGGCTTTTAAAGAAATATCTTTTCCCAATTGCAATTCATTA 792
Db 115 -----AlaLeuSerPheSerGlyPheSerSerLeu 124
QY 793 CTT----- 795
Db 125 LeuIleAspSerAlaProAlaThrGlyThrSerGlyGlyLysGlyAlaIleCysValThr 144
QY 795 ----- 795
Db 145 AsnThrGluGlyGlyThrAlaThrPheThrAspAsnAlaSerValThrLeuGlnLysAsn 164
QY 795 ----- 795
Db 165 ThrSerGluLysAspGlyAlaAlaValSerAlaTyrSerIleAspLeuAlaLysThrThr 184
QY 796 -----GCCGTACTGCCCTGCT 810
Db 185 ThrAlaAlaLeuLeuAspGlnAsnThrSerThrLysAsnGlyGlyAlaLeuCysSerThr 204
QY 811 GCAACGACT-----AATAAGGGTAGCCAGACT---CCGACGACAAACATCTACA 855
Db 205 AlaAsnThrThrValGlnGlnAsnSerGlyThrValThrPheSerSerAsnThrAlaThr 224
QY 856 CCGTCTAATGGTACTATTATTCT- 879
```

```
Db 225 AspLysGlyGlyGlyIleTyrSerLysGluLysAspSerThrLeuAspAlaAsnThrGly 244
QY 879 ----- 879
Db 245 ValValThrPheLysSerAsnThrAlaLysThrGlyGlyAlaTrrpSerSerAspAspAsn 264
QY 879 ----- 879
Db 265 LeuAlaLeuThrGlyAsnThrGlnValLeuPheGlnGluAsnLysThrThrGlySerAla 284
QY 879 ----- 879
Db 285 AlaGlnAlaAsnAsnProGluGlyCysGlyGlyAlaIleCysCysTyLeuAlaThrAla 304
QY 880 -----AAAAACAGATCTTTTGTACTCAATATAGAAAGTTCTCTCATTTCTATAGTAAATTTA 933
Db 305 ThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMetSerPheThrSerAsnThr 324
QY 934 GTCTCTGGAGATGGGGAGCTATAGATCTTAAGAGCTTAACGGTTTCAAGGAATTAGCAAG 993
Db 325 ThrThrAlaAsnGlyGlyAlaIleTyrAlaThrLysCysThrLeuAspGlyAsnThrThr 344
QY 994 CTTTGTGCTTCCAAAGAAAATACTCTCAAGCTGAT---GGGGAGCTTGTCAAGTAGTC 1050
Db 345 Leu---ThrPheAspGlnAsnThrAlaThrAlaGlyCysGlyGlyAlaIleTyrThrGlu 363
QY 1051 ACC---AGTTCTCTGCTATGCTAACGAGGCTCTTATGCTTGTAGCGAATGTTGCA 1107
Db 364 ThrGluAspPheSerLeuLysGlySerThrGlyThrValThrPheSerThrAsnThrAla 383
QY 1108 CGAGTAAAGAGGGGGGAGGATT----- 1128
Db 384 LysThr---GlyGlyAlaLeuTyrSerLysGlyAsnSerSerLeuThrGlyAsnThrAsn 402
QY 1129 -----GCTGCTGTTCCAGGATGGG 1146
Db 403 LeuLeuPheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAsnGlnGluGly 422
QY 1147 CAGCAGGGAGTG-----TCATCATCTACTTCAACA----- 1176
Db 423 CysGlyGlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysLysGlyLeu 442
QY 1177 -----GAAGATCCAGTAGTAAGTTTTCAGAAATACTGCG----- 1212
Db 443 TrpIleGluAspAsnGluAsnValSerLeuSerGlyAsnThrAlaThrValSerGlyGly 462
QY 1213 -----GTAGAGTTTGTGATGGGAAC 1230
Db 463 AlaIleTyrAlaThrLysCysAlaLeuHisGlyAsnThrThrLeuThrPheAspGlyAsn 482
QY 1231 GTAGCCCGA---GTAGGAGGAGGATTAC----- 1257
Db 483 ThrAlaGluThrAlaGlyGlyAlaIleTyrThrGluThrGluAspPheThrLeuThrGly 502
QY 1258 TCTACGCGGAACCTGCTTTTCTGTAATAAT----- 1287
Db 503 SerThrGlyThrValThrPheSerThrAsnThrAlaLysThrAlaGlyAlaLeuHisThr 522
QY 1288 ---GGAAAAACCTGTTTCTCAACAATGTTGCTTCTCCTGTTTACATTGCTGTACAGCAA 1344
Db 523 LysGlyAsnThrSerPheThrLysAsnLysAla-----LeuValPheSerGlyAsn 539
QY 1345 CCAAAGTGGACAGGCTTCTAATACAGTAATAATTACGAGATCGAGGAGCTATCTTCTTC 1404
Db 540 SerAlaThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyGlyAlaIleLeu 559
QY 1405 TGT-----AAGATGTCGCG 1419
Db 560 CysAsnIleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAsnGluSer 579
QY 1420 CAAGCAGGATCCCAATTAACCTCTGGATCAGTTTCC----- 1452
```


580 LeuSerPheIleAenAsnThrAlaLysArgSerGlyGlyGlyIleTyrAlaProLysCys 599
1453 ---TTTGATGAGAGGGAGTAGTTTCTTTAGTAGCAATGATGCTGGGAAAGGGGA 1509
600 ValIleSerGlySerGluSerIleAenPheAepGlyAenThrAlaGluThrSerGlyGly 619
1510 GCTATTTATGCCAAAAGCTCTCGTTGCTAACTGCTGGCCCTGTACAAATTTTAAAGGAT 1569
620 AlaIleTyrSerLysAenLeuSerIleThrAlaAenGlyProValSerPheThrAenAen 639
1570 ATCGTAATGATGGTGAGCGCATTTATTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCT 1629
640 SerGlyGlyLysGlyAlaIleTyrIleAlaAepSerGlyGluSerLeuGluAla 659
1630 GATTATGAGATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAGAATGCTGCC 1689
660 IleAepGlyAepIleThrPheSerGlyAen-----ArgAlaThrGluGlyThrSer 676
1690 GATGTTAATGGGTAACTGTGCTCACAGCCATTTGATGGGATCGGAGCGGAGGAATA 1749
677 ThrProAen-----SerIleHisLeuGlyAlaGlyIle 689
1750 ACGCATTAAAGAGCTAAAGCGGCATCAGATTCTCTTTAATGATCCCATCGAGATG--- 1806
690 ThrLysLeuAlaIleAalProGlyHisThrIleTyrPheTyrAepProIleThrMetGlu 709
1807 -----GCAACGGGA----- 1815
710 AlaProAlaSerGlyGlyThrIleGluGluLeuValIleAenProValValLysAlaIle 729
1816 -----AATAACGCGCGAGCGAGTCTTCAAACTCTCAAAATTAAC 1857
730 ValProProGlnProLysAenGlyProIleAlaSerValProValValProValAla 749
1858 GATGTGAAGCATACACAGGGGATATCTTTT----- 1890
750 ProAlaAenProAenThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAep 769
1891 -----GCTAATGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGCA 1938
770 AlaSerIleProAlaAenThrThrThrIleLeuAenGlnLysIleAenLeuAlaGlyGly 789
1939 AGGATTGTTCTCGTAAAGCGAAATATATCAGTAAATCTCTAAGTCACAGAGGTGG 1998
790 AenValValLeuLysGluGlyAlaThrLeuGlnValTyrSerPheThrGlnGlnProAep 809
1999 AGT---CTGTATATGAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACAA 2055
810 SerThrValPheMetAalAalGlyThrThrLeuGluThrThrThrAen----- 826
2056 CAGCTCTCGCGCTAATCAGTTGATCAGCTTTTCCAACTGCAATTTGTCTCTTCTCT 2115
827 -----AenThrAepGlySerIleAepLeuLysAenLeuSerValAenLeuAalA 843
2116 TTGTAGCAAAATGAGTTACGATCCTCTACCAATCTCTCCAGCCCAAGATTTCTCAT 2175
844 LeuAepGlyLysArgMetIleThr----- 851
2176 CTGCAGTCATTGGTAGCACTGCTGTTTCAATATTAGTGGGCTATCTTTT 2235
852 -----IleAlaValAenSerThrSerGlyGlyLeuLysIleSerGlyAepLeuLysPhe 869
2236 GAGGATTGGATGATCAGCTTATGATAGGTATGATGGTCTTAATCAAAAATC 2295
870 HisAenAenGluGlySerPheTyrAepAen-----ProGlyLeuLysAlaAenLeu 886
2296 AATGCTCTGAATATACAGTTA---GGGACTAAGCCCCCAGCTAAT----- 2337
887 AenLeuProPheLeuAepLysSerSerThrSerGlyThrValAenLeuAepPheAen 906
2338 ---GCCCATCAGATTTCACCTAGGGAATGAGATGCTTAAGTATGGTATCAGGAGGC 2394
907 ProIleProSerSerMetAla-----AlaProAepTyrGlyThrGlnGlySer 922

2395 TGGAGCTTGGTGGGATCTCTAATAACAGCAAAATAATGGTCTTATATCTCTCAAGAGCTACA 2454
923 TrpThrLeuVal-----ProLysValGlyAlaGlyLysValThrLeuValAlaGlu 940
2455 TGGACTAAACTGGGTATATCCCTGGGCTGAGCGAGTAGCTTCTTTGGTTCCTCAATAGT 2514
941 TrpGlnAlaLeuGlyTyrThrProLysProGluLeuArgAlaThrLeuValProAenSer 960
2515 TTATGGGATCCATTTTAGATATACGATCTGGCATTCAGCAATTCAGCAAGGTGTGGAT 2574
961 LeuTrpAenAlaTyrValAenIleHisSerIleGlnGluIleAlaThrAlaMetSer 980
2575 GGGCGCTCTTATTCGAGGATTAATGGTTTCTCGAGTTTCGAATTTCTCTATCATGAC 2634
981 AspAlaProSerHisProGlyIleTrpIleGlyGlyAenAlaPheHisGlnAep 1000
2635 CGCGATGCTTTAGTACGGATATCGGTATATTTAGTGGGGTTATCTCTTAGGAGCAAC 2694
1001 LysGlnLysGluAenAlaGlyPheArgLeuIleSerArgGlyTyrIleValGlyGlySer 1020
2695 TCCTACTTTGGATCATCG---ATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGACT 2751
1021 MetThrThrProGlnGluTyrThrPheAlaValAlaPheSerGlnLeuPheGlyLysSer 1040
2752 AAAGATTATGATGTGCTGCTTCCAATCATCATGCTTGCATAGGATCCGTTTATCTATCT 2811
1041 LysAepTyrValValSerAepIleLysSerGlnValTyrAlaGlySer-----LeuCys 1058
2812 ACCCAACAGCTTTATGGATCTCTATTTGTTCCGAGATGGGTTTATCCGTGCT----- 2865
1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgHisValLeuSerLys 1078
2866 -----AGCTACGG 2873
1078 sValLeuProGluLeuProGlyGluThrProLeuValLeuHisGlyGlnValSerThrGly 1098
2874 GTTTGGGAATCAGCATATGAAACCTCATATACATTTTCAGAGAGGAGCGATGTTGTTG 2933
1098 YArgAenHisHisAenMetThrThrLysLeuAlaAenAenThrGlnGlyLysSerAepTyr 1118
2934 GGATAATAACTGCTGCTGGAGATTTGGAGCGGATTTACCGATTGATTTATCTCCATC 2993
1118 pAepSerHisSerPheAlaValGluValGlySerLeuProValAepLeu-----As 1136
2994 TAAGCTCTATTGAATCAGTTGCTGCTTCTGCAAGCTGAGTGTCTTATGCGCATCA 3053
1136 nTyrArgTyrLeuThrSerTyrSerProTyrValLysLeuGlnValValSerValAenGly 1156
3054 TGAATCTTTTACAGAGAGCGGATCAAGCTCGGCGATTTCAAGAGCGGACATCTCTTAA 3113
1156 nLysGlyPheGlnGluValAlaAalAepProArgIlePheAepAlaSerHisLeuValA 1176
3114 TCTATCAGTTCTGTTGGAGTGAAGTTTGATGCTGATTTCTAGTACACATCTTAATAATA 3173
1176 nValSerIleProMetGlyLeuThrPheLysHisGluSerAlaLysProSerAlaLe 1196
3174 TAGCTTTATGGCGCTTATATCTGATGCTTATTCGCACCATCTCTGCTACTGAGACAAC 3233
1196 uLeuLeuThrLeuGlyTyrAlaValAalAalAalAalAalAalAalAalAalAalAalA 1216
3234 GCTCTATCCCATCAAGAGACATGACAGATGCTCTTTTATTTAGCAAGATGAGT 3293
1216 r---LeuThrAenGlyThrSerThrPheAlaThrAenLeuSerArgGlnAlaPhe 1235
3294 TGTGTTAGGAGTCTATGCTCTCTTAAAGTAATATAGAAAGTATATAGGATGATG 3353
1235 ePheAlaGluAalSerGlyHisLeuLysLeuHisGlyLeuAalAalAalAalAalAalA 1255
3354 AAGATATGATGATTCAGATGCTCTCGAGGCTATGTTTGGTTCAGTCAGAGTAGGTCG 3413
1255 ySerCysGluLeuArgSerSerSerArgSerTyrAenAlaAenCysGlyThrArgTyrSe 1275

```

QY      3414 GTTC 3417
Db      1275 rPhe 1276

RESULT 29
Q823X0 CHLCV
ID      Q823X0_CHLCV PRELIMINARY; PRT; 866 AA.
AC      Q823X0_
DT      01-JUN-2003 (TREMBLrel. 24, Created)
DT      01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Polymorphic outer membrane protein G family
DE      protein/autotransporter.
DS
GN      OrderedLocusNames=CCAO0283;
OS      Chlamydomophila caviae.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX      NCBI_TaxID=83557;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=GPIC;
RX      MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA      Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA      Heidelberg J.F., Holtzaple E.K., Khouri H.M., Federova N.B.,
RA      Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA      White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA      Bavoli P.M., Fraser C.M.;
RT      "Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
RT      examining the role of niche-specific genes in the evolution of the
RT      Chlamydiaceae."
RL      Nucleic Acids Res. 31:2134-2147(2003).
DR      EMBL; AA016995; AAP05034.1; -; Genomic_DNA.
DR      TIGR; CCA00283; -.
DR      GO; GO:0019867; C:outer membrane; IEA.
DR      InterPro; IPR006315; Autotransporter.
DR      InterPro; IPR005546; Auto transptbeta.
DR      InterPro; IPR011427; ChlamPMP M.
DR      InterPro; IPR003368; Chlamydia_PMP.
DR      Pfam; PF03797; Autotransporter; 1.
DR      Pfam; PF0548; ChlamPMP M; 1.
DR      Pfam; PF02415; Chlam PMP; 5.
DR      TIGRFAMS; TIGR01414; autotrans bar1; 1.
DR      TIGRFAMS; TIGR01376; POMP_repeat; 4.
DR      Complete proteome.
SQ      SEQUENCE 866 AA; 92969 MW; 6CFA14937C69DB30 CRC64;

Alignment Scores:
Pred. No.:      3 92e-58      Length:      866
Score:          1008.00      Matches:    303
Percent Similarity: 41.7%      Conservative: 144
Best Local Similarity: 28.2%      Mismatches: 358
Query Match:      12.8%      Indels:    268
DB:              2          Gaps:      30

US-10-701-844-1 (1-4435) x Q823X0_CHLCV (1-866)
QY      382 ATGCAACGCTTTCCATAAGTTCTTTTCAATGATNTAGCTTATTTCTGCTGCTCT 441
Db      1 MetLysHisProValTyrTrpPheLeuValSerSerGlyLeuLeuAlaSer---ThrSer 19
QY      442 TTAATCGGGGGGATGACGACGAAATCATGGTTCTCAAGGAATTTACGATGGGAG 501
Db      20 LeuSerPheAlaAlaValGlnGluThrLeuAsnSerSerAspSerTyrAsnGlyAsn 39
QY      502 ACGTTAACTGTATCATTT-----CCCTATACTGTTATA 534
Db      40 AlaThrThrAspAlaPheGlnThrLysGluThrThrGlyAlaGluTyrThrCysGlu 59
QY      535 GGAGATCCGAGTGGGACTACTGTTTTTTCTGACGAGAGTTAAACATTAATAAATCTTGAC 594
Db      60 GlyAsn-----ValCysIleThrTyrAlaGlyLysThr 71
QY      595 AATCTATGACGCTTTCCTTTAAGTTGTTTGGGAACCTATTATAGGAGTTTACTGTT 654
Db      279 AsnLysIleIleThrThrGlyThrProSerThrLysArgAsn----- 292

102 -----ProLeuThrLysSerCysPheThrGluThrThrGluAsnLeuThrPhe 87
655 TTAGGAGAGACACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACAAATGGGCGACT 714
88 IleGlyGlnGlyTyrSerLeuCysPheAspAsnIleThrAlaThrAlaLysProAlaAla 107
715 CTAAGTAATAGCGCTGCTGAGACTGTTTACTATTAGGGTTTTTAAAGAATTATCCCTTT 774
108 IleGluValSerAlaAlaAspLysThrLeuSerIleSerGlyPheSerLeuPheSerCys 127
775 TCCAAATTCGAATTCATTACTTCCGCTACTGCTGCAACGACTATAAAGGGTACCCAG 834
128 SerTyrCys-----ProProGlyThrThrGlyGlnGlyAlaIle 140
835 ACTCCGACGACAAACATCTACACCGTCTAATGCTACTATTATTCTTAAACACAGATCTTTG 894
141 LysSerGlyGlyThrAlaThr----- 147
895 TTACTCAATAATAGAGAGTTCTCATTTCTATAGTAATTTAGTCTCTGGAGATGGGGAGCT 954
148 PheAspAsnAsnSerLysLeuPheLysSerAsnCysSerSerGluGluGlyAla 167
955 ATAGATGCTAAGAGCTTAACGGTTCAAGGAATTACCAAGCTTTGTCTTCCCAAGAAAT 1014
168 IleAsnCysLysSerLeuThrLeuLysAsn---SerSerValCys----- 181
1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCAAGTTCTCTGCTATGGCTAAC 1074
181 ----- 181
1075 GAGGCTCCTATTGCTTTTGTAGCGAATGTTGAGGAGTAAGAGGGGAGGGATGCTGCT 1134
181 ----- 181
1135 GTTCAGGATGGGACAGGAGGTGTCATCATCTTCAACAGAGATCCAGTAGTAAGT 1194
182 -----AlaAsn 183
1195 TTTTCCAGAAATACTCGGTAGAGTTTGATGGACGTAGCCGAGTAGGAGGGGAT 1254
184 PheIleLysAsnSerSerAspLys----- 191
1255 TACTCTACGGGAACGTTGCTTTCTCTGAATAATGGAATAACCTTTTCTCAACAATGT 1314
191 ----- 191
1315 GCTTCTCTGTTTACATTCCTGCTAGCAACCAACAGTGGACAGGCTTCTAATACGAGT 1374
191 ----- 191
1375 AATAATTACGAGATGGAGAGCTATCTTCTGTAAGAATGGTGGCAAGCAGGATCCAAT 1434
192 -----LysGlyGlyAlaIleTyrCys----- 198
1435 AACTCTGGATCAGTTCTTCTGATGGAGAGGAGTAGTTTCTTTTAGTACCAATGAGCT 1494
199 SerAspGluAsnLeuGlnLeuGluAsnAsnAspGlnMetLeuPheSerGluAsnThrSer 218
1495 GCTGGGAAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTGCTAAGTGGCCCTGTGA 1554
219 LysGluGluGlyAlaIleTyrAlaLysLysLeuSerIleIleSerGlyGlyProThr 238
1555 CAATTTTAAAGGAATATCGCTAAT-----GATCGTGGAGGAGTTTATTTA 1599
239 LeuPheSerAsnAsnSerThrSerLysAlaAlaAspProLysGlyAlaIleCysIle 258
1600 GGAGATCT---GGAGAGCTCAGTTTATCTGCTGATTATATGAGATATATTTTCGATGG 1656
259 AlaAspAlaAspSerGluCysSerLeuThrAlaGluAsnGlyAspIleIlePheAspGly 278
1657 AATCTTAAAGAACACAGCCAAAGAGATGTCGCGCATGTTTAATGGCGTAACCTGTGCTCA 1716
279 AsnLysIleIleThrThrGlyThrProSerThrLysArgAsn----- 292

```

```
QY 1717 CAAGCCATTTCGATGGGATCGGAGGGAATAAAGACATTAAGACGTAAAGCAGGGCAT 1776
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 ---SerIleAspLeuGlySerGlyGlyLysPheSerGlnLeuArgAlaAraAspGlyPhe 311
QY 1777 CAGATTCTCTTAATGATCCCATCGAGATGCAAGGAAATACACCGCAGCGAGTCT 1836
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 GlyValPhePheTyraAspProile-----AlaAsnAsn-----GlySerAsp 325
QY 1837 TCCAACTCTTAAATTAACGATGGTGAAGGA-----TACACAGGGGATATTGTT 1887
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 ThrAspThrLeuGluIleAsnLysAlaAspGlyAlaAlaThrTyraSerGlyArgIleVal 345
QY 1888 TTGCT-----AATGGAAGCAGTACT 1908
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 PheSerGlyGlyLysLeuThrGluAspGlyLysGlnValThrAspAsnLeuLysSerPhe 365
QY 1909 TTGTACCAAAATGTTACGATAGACGAAGGATTGTTCTTCGTGAAGGCAAAATTA 1968
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366 PheLysGlnProLeuThrValGlySerGlySerPheValLeuLysAsnGlyValThrVal 385
QY 1969 TCAGTGAATCTTAAGTCAGACAGGTGGAGTCTGTATATGGAAGCTGGGAGTACATGG 2028
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 SerAlaLysGlnIleThrGlnSerGlyGlyAlaIleGluMetAspAlaGlyThrAsn--- 404
QY 2029 GATTGTTAACTCCAAACCAACCAACAGCCTCTCGCGCTAATCATGTTGATCAGCTT 2088
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
405 -----LeuThrSer 407
QY 2089 TCCAATCGCATTTGCTCTTCTCTTTGTAGCAACATGACGATTACGAATCCTCCT 2148
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
408 ThrThrGluAspIleSerLeuSerAsnLeuValIleAsnThrAlaSerLeuGlyGlyGly 427
QY 2149 ACCAATCTCAGCGCAAGATTCTCATCTCGAGTCATTGTTAGCACAACTGCTGTTCT 2208
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
428 GlyValProLeuAlaAlaGln-----IleSerAlaGluGlyThrAsnLysSer 443
QY 2209 GTTACAATTAGGGCCATCTCTTTTGGAGATTTGGATGATACAGCTTATGATAGGTAT 2268
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
444 ValThrIleSerSer---LeuAsnLeuValAspAlaAspGlyAsnGlyTyraGluTyraPro 462
QY 2269 GATTGGCTAGTCTTAATCAAAAATCAATCTCCGAAATTCACAGTTA---GGCAGCTAAG 2325
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 ValPheSerThrThrArgGlyPheProSerIleIleGluAlaLysAlaAsnGlyThrGly 482
QY 2326 CCCCAGCTAATGCCCATCAGATTGACTCTAGGGAATGAGATCCCTAAGTATGGCTAT 2385
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
483 ThrProThrIleProThrThrHisLeuThrAspHisAlaProAlaAlaHisTyraGlyTyra 502
QY 2386 CAAGAGAGCTGGAAGCTTGGCTGGATCCTTAATACAGCAATAATGTCCTTATCTCTG 2445
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
503 GlnGlyLeuTrpThrThrSerTrpAlaGlnGlyThrAlaThrThrSerGlnLeuAlaThr 522
QY 2446 AAAGCTACATCGACTAAACCTGGGTATATCTCGGCTGAGCGAGTACTTCTTGGTT 2505
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 LeuAla---TrpGlnGlnThrGlyTyraAsnProAsnProGluLysGlnGlyProLeuVal 541
QY 2506 CCAATAGTTATGGGATCCATTTAGATATACATCTCGGCTATCAGCAATTAACAGCA 2565
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
542 ProAsnThrLeuTrpGlySerPheSerAspValArgAlaIleGlnAsnLeuMetAspIle 561
QY 2566 AGTGTGATGGGCGCTCTTATATGTCGAGATATATGGTTCTCGAGTTTCGAATTTCTTC 2625
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
562 SerValAsnGlyAlaAspTyraGlnArgGlyLeuTrpAlaSerGlyLeuAlaAsnPheLeu 581
QY 2626 TATCATGACCGGATCTTTAGTTCAGGATATCGGTATATATAGTGGGGTTATCTCTTA 2685
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
582 GlnLysSerGlyThrGluThrLysArgLysPheArgHisIleSerAlaGlyTyraValLeu 601
QY 2686 GGAGCAAACTCC---TACTTTGGATCATCGATGTTTGGTCTAGATATTTACCGAAGTATTT 2742
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
602 GlyAlaTyraAlaLysThrLeuSerAspAspValPheSerAlaAlaPheCysGlnLeuPhe 621
```

```
QY 2743 GGTAGATCTAAAGATTATGCTGTGTCTTCCATCATCATCTGTCATAGGATCGTT 2802
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
622 GlyArgAspLysAspTyraLeuValSerLysAsnAsnSerAsnIleTyraAlaGlySerIle 641
QY 2803 TATCTATCTACCCCAACAAGCTTTATGTGGATCTTATTTGTCGAGATGCGTT--- 2856
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
642 TyrTyraGlnHisThrSer-----PheTrpAspAlaTrpAspAsn 654
QY 2857 -----ATCCGTGCTAGCTAC 2871
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
655 LeuLeuGlnSerThrLeuGlyAlaGlnAlaProLeuValLeuAsnAlaGlnLeuThrTyra 674
QY 2872 GGGTTTGGGAATCAGCATATGAAAACCTCATATACA-----TTTGCAGAGGAGCGAT 2925
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
675 SerHisThrSerAsnAspMetLysThrAsnMetThrTyraSerTyraAlaProGlnGlyVal 694
QY 2926 GTT-----CGTTGGATAATAAATCTGCTGGCTGGAGAGATTGGAGCG 2967
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
695 ValTyraProGluIleLysGlyAspTrpGlyAsnAspCysPheGlyValGluLeuGlyAla 714
QY 2968 GGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGAATGATGTTGCGTCTTCTG 3027
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
715 ThrValProIleGluSerProTyraSerSerLeuPhe---AspMetTyraSerProPheLeu 733
QY 3028 CAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGGCGAT---CAAGCT 3084
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
734 ArgPheGlnLeuValTyraAlaHisGlnGluAspPheLysGluAsnAsnSerThrGluGly 753
QY 3085 CGGGCATTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTTGAT 3144
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
754 ArgTyraPheGluSerSerAspLeuThrAsnLeuSerMetProIleGlyValLysPheGlu 773
QY 3145 CGATTGTTCTAGTACACATCTCTAATAATATAGTCTTATGCGGCTTATATCTGTGATGCT 3204
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
774 ArgPheSerAspAsnAspIleAlaSerTyraAsnValThrLeuAlaTyraAlaProAspLeu 793
QY 3205 TATCGCACCATCTCTGTTAGCAAGACGCTCTATCCCATCAAGAGACA-----TGG 3258
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
794 ValArgSerAsnProAspCysLysThrSerLeuLeuValSerProThrThrAlaValTrp 813
QY 3259 ACAACAGATCGCTTCTTATGCAAGACATGAGTGTGTTAGGAGTCTATCTATGCT 3318
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
814 LeuThrLysAlaThrAsnLeuAlaArgHisAlaPheIleValLysAlaGlyAsnTyraLeu 833
QY 3319 TCTTAAACAAGTAAATAGAAATATATGCGCATGGAAGATATAGATATCGAGATGCTTCT 3378
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
834 SerLeuSerSerAsnPheGluIlePheSerGlnPheGlyPheGluLeuArgGlySerSer 853
QY 3379 CGAGGCTATGTTGAGTCAGGAGTAGAGTCCGGTTC 3417
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854 ArgThrTyraAsnValAspLeuGlySerLysIleGlnPhe 866

RESULT 30
Q9RB71.CHLPN PRELIMINARY; PRT; 772 AA.
ID Q9RB71; Q7BX22;
AC Q9RB71; Q7BX22;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Pmp_3 (Outer membrane protein 5)
GN Name=pmp_3_2; OrderedLocuNames=CpB0018;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OC NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=J138;
RX MEDLINE=2033049; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
```

RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=FW-183;
 RA Geng M.M., Schuhmacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,
 RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.,
 RT "The genome sequence of Chlamydia pneumoniae TW183 and a comparison with
 RT other Chlamydia strains based on whole genome sequence analysis.";
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BA000008; BAA98226.1; -; Genomic DNA.
 DR EMBL; AB017157; AAP97951.1; -; Genomic DNA.
 DR PIR; H86492; H86492.
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Auto transportbeta.
 DR InterPro; IPR011427; Chlamydia_PMP.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07548; Chlamydia_PMP; 1.
 DR Pfam; PF02415; Chlamydia_PMP; 6.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
 SQ SEQUENCE 772 AA; 82931 MW; 484FC56D635801EB CRC64;

 Alignment Scores:
 Pred. No.: 2,96e-56 Length: 772
 Score: 979.50 Matches: 261
 Percent Similarity: 46.2% Conservative: 144
 Best Local Similarity: 29.8% Mismatches: 323
 Query Match: 12.4% Indels: 149
 DB: 2 Gaps: 24

 US-10-701-844-1 (1-4435) x Q9RB71_CHLPN (1-772)
 QY 919 TTCTATAGTAATTAAGTCTCTGGAGATGGGGAGCTATAGATCTTAAGCTTAACGGTT 978
 DB 3 PheSer-LysAsn-PheSer-ThrAspAsnGlyGlyAlaIleThrAlaLysThrLeuSerLeu 22
 QY 979 CAAGGAATTAGCAAGCTTGTCTTCCAGGAATTAAGTCTCAAGCTGATGGGGAGCT 1038
 DB 23 ThrGlyThrMetSerAlaLeuPheSerGluAsnThrSerSerLysGlyAla 42
 QY 1039 TGTCAAGTAGTACCAGTTTCTCTGCTATGCTTAACGAGGCTCTATTGCTTTGTAGCG 1098
 DB 43 IleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPhe 60
 QY 1099 AATGTTGCAGGAGTAGAGGGAGGAGTGTCTGCTTCCAGGATGGGACGAGGAGTG 1158
 DB 60 ----- 60
 QY 1159 TCATCATCTACTCAACAGAGATCCAGTAGTAGTATGTTTCCAGAAATACGCGGTAGAG 1218
 DB 61 SerAspAsnThrSerSerAspSer----- 68
 QY 1219 TTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTTC 1278
 DB 69 -----GlyAlaAlaIlePheThrGluAlaSerValThrIle 80
 QY 1279 CTGAATATGGAATAACCTTTGTTCTCAACAATGTTGCTTCTCTGTTTACATGCTGCT 1338
 DB 81 SerAsnAsnAlaLysValSerPheIleAspAsn----- 91
 QY 1339 AAGCAACAACAGTGGACGAGCTTCTTAATACGAGTAGTAATATACGAGATGGAGGAGCT 1398
 DB 92 -----LysValThrGlyAlaSerSerSerThrThrGlyAspMet-----SerGlyGlyAla 108
 QY 1399 ATCTTCTGTAAGATGTCGCCAGCAGGATCCAACTAAGTCTGATCAGTTTCTTGTGAT 1458
 DB 109 IleCysAlaTyrls-----ThrSerThrAspThrLysValThrLeuThr 123
 QY 1459 GGAGAGGGAGTAGTATTTCTTTAGTAGCAATAGCTGCTGGGAAAGGGAGGAGCTATTAT 1518
 DB 124 GlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrThrAlaGlyAlaIleTyrl 143

QY 1519 GCCAAAAGCTCTCGTGTCTAATCTGCGCCTGTACAAATTTTAAAGAAATATCCCTAAT 1578
 DB 144 VallysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerValaAsn 163
 QY 1579 -----GATGCTGGAGCATTTATTTAGGAGATCTGGAGAGCTCGATTTA 1623
 DB 164 GlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGluLeuSerLeu 183
 QY 1624 TCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAGAGAT 1683
 DB 184 SerAlaAspSerGlyAspIleValPheLeuGlyAsn-----ThrValThrSerThr 200
 QY 1684 GCTGCCGATTAATGGGCTAACTGTCTCTCACAGCCATTTCCATGGGATCGGAGGG 1743
 DB 201 ThrProGlyThrAsn-----ArgSerSerIleAspLeuGlyThrSerAla 215
 QY 1744 AAAATAACGACATTAAGACTAAACAGCGCATCAGATTCTCTTTAATCATCCATCGAG 1803
 DB 216 LysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleTyrlPheTyrlAspProIleThr 235
 QY 1804 ATGCCAAACGAAATAAACCCAGCGCAGTCTCTCCAACTTTCTAAAAATTAAACGAGGT 1863
 DB 236 ThrGlySerSerThrThr-----ValThrAspValLeuLysValaAsnGluThr 251
 QY 1864 GAAGGA-----TACACAGGGATATTGTTTTGCT----- 1893
 DB 252 ProAlaAspSerAlaLeuGlnTyrlThrGlyAsnIleIlePheThrGlyGlyLysLeuSer 271
 QY 1894 -----RATGGAACGACTACTTTGTACCAAAATGTTACGATA 1929
 DB 272 GluThrGluAlaAspSerLysAsnLeuThrSerLysLeuLeuGlnProValThrLeu 291
 QY 1930 GAGCAAGGAAGGATTTCTCTGTAAGGCAAAATTAATCATGTAATCTCTTAAGTCAG 1989
 DB 292 SerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThrGlnAlaPheThrGln 311
 QY 1990 ACAGGTGGAGT---CTGTATATGGAAGCTGGAGTACATGGAGATTTTGTAACTCCACAA 2046
 DB 312 GlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu----- 326
 QY 2047 CCACCAACAGCCTCTCTCGCGCTAATCATGTTGATCAGCTTTTCCATTCGATTTGCT 2106
 DB 327 -----ProAlaAspThrSer-----ThrIleAsnAsnLeuValIleAsn 339
 QY 2107 CTTTCTCTTTGTTAGCAAAACATGCAATGCAATCTCCTACCAATCTCCAGCGCAA 2166
 DB 340 IleSerSerIle-----Asp 344
 QY 2167 GATTCTCATCTGCACTCATTTGGTAGC---ACAACTGCTGGTCTGTTCATTAAGTAGGG 2223
 DB 345 GlyAlaLysLysAlaLysIleGluThrLysAlaThrSerLysAsnLeuThrLeuSerGly 364
 QY 2224 CCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGTATGATGATGCTGCTTCT 2283
 DB 365 ThrIleThrLeuLeuAspProThrGlyThrPheTyrlGluAsnHisSer---LeuArgAsn 383
 QY 2284 AATCAAAAATCAATCTCTGTAATTAACAGTTA---GGGACTAAAGCCCCAGCTAATGCC 2340
 DB 384 ProGlnSerTyrlAspIleLeuGluLeuLysAlaSerGlyThrValThrSerThrAlaVal 403
 QY 2341 CCATCAGATTGACTCTAGGGAATGAGATGCCCTTAAGTATGCTATCAAGAAAGCTGG--- 2397
 DB 404 ThrProAspProIleMetGlyGlyLysPhe---HisTyrlGlyTyrlGlnGlyThrTrpGly 422
 QY 2398 AAGCTTGGCGGATCTTAATACAGCAATAATATGCTCTTATCTCTGAAGAGCTACATGG 2457
 DB 423 ProIleValTrpGlyThrGlyAlaSerThrThrAlaThrPhe-----AsnTrp 438
 QY 2458 ACTAAACTGGGTATTAATCTCGGCGCTGAGGAGTAGCTTCTTTGGTTCCAAATAGTTTA 2517
 DB 439 ThrLysThrGlyTyrlProAsnProGluArgIleGlySerLeuValProAsnSerLeu 458
 QY 2518 TGGGGATCCATTTTATAGATATACGATCTGCGCATTTACGAAATTCAGCAATTCAGCAAGTGTGGATGGG 2577

DR TIGRFPAMs; TIGR01376; POMP_repeat; 5.
 KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
 FT SIGNAL 1 26 Potential.
 FT CHAIN 27 922 Probable outer membrane protein pmp1.
 FT CONFLICT 14 14 F -> L (in Ref. 1).
 FT CONFLICT 251 251 A -> P (in Ref. 5).
 FT CONFLICT 375 375 Y -> C (in Ref. 1).
 FT CONFLICT 606 606 D -> N (in Ref. 1).
 FT CONFLICT 836 836 S -> P (in Ref. 5).
 SQ SEQUENCE 922 AA; 100458 MW; DFF2AB6333AB031C CRC64;

Alignment Scores:
 Pred. No.: 8 41e-56 Length: 922
 Score: 973.00 Matches: 306
 Percent Similarity: 41.5% Conservatives: 148
 Best Local Similarity: 28.0% Mismatches: 386
 Query Match: 12.3% Indels: 253
 DB: 1 Gaps: 35

US-10-701-844-1 (1-4435) x PMP1_CHLPN (1-922)

QY 375 TTGTGTGATGCAACGCTTTTCCATAAGTCTTTCTTTCAAT-----GAT 419
 |||||
 Db 5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21
 |||||
 QY 420 TCTAGCTTATCTTGTCTCTCTTTAAATCGGGGGGATATGCGACGAATCATGTTCC 479
 |||||
 Db 22 ThrSerLeu-SerAlaThrIleSer-----LeuThrPr 33
 |||||
 QY 480 TCANGAAATTTAGATGGGGAGAGCTTAACGTGTATCATTTCTCTACTCTTATAGGAGA 539
 |||||
 Db 33 OGUAspSerPheHisGlyAspSerGlnAsnAlaGUAspSerTyAsnVal-----50
 |||||
 QY 540 TCGAGTGGGACTACTGTTTTTTCTCAGGAGAGTTTAAACATTAATAAATCTTGACAATTC 599
 |||||
 Db 51 -GlnAlaGlyAspValTy-SerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70
 |||||
 QY 600 TATTGAGCTTTGCCCTTTAGTTGTTTGGAACTATTATAGGAGTTTACTGTTTAGG 659
 |||||
 Db 70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGl 88
 |||||
 QY 660 GAGAGGACTCGTGCACCTTCAGACATACGAGCTCTTACA-----AATGGGCAGC 713
 |||||
 Db 88 yAsnHisGlyLeuTy-PheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 108
 |||||
 QY 714 TCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTAAAGAAATATCCTT 773
 |||||
 Db 108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128
 |||||
 QY 774 TTCCAATTGCAATTCAATTACTTCGCTGATGCTGCTGCAACGACTAATAAGGTTAGCCA 833
 |||||
 Db 128 e-----lLeG1 130
 |||||
 QY 834 GACTCCGACGACAACATCTACCGCTAATGCTACTATTATTATCTAAACAGATCTTTT 893
 |||||
 Db 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTySerLysAsnAlaLeuMe 148
 |||||
 QY 894 GTTACTCAATAATGAGAAGTTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGC 953
 |||||
 Db 148 tLeuLeuAsnAsnTyValValArgPheGluGlnAsnGlnSerLysThrLysGlyAl 168
 |||||
 QY 954 TATAGATGCTAAGAGCTTAACGTTTCAAGAAATTTAGCAAGCTTTGTCTTCCAGNAAA 1013
 |||||
 Db 168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyAspSerValSerPheTyGlnAs 188
 |||||
 QY 1014 TACTGCTCAAGCTGATCGGGGAGCTTCTCAAGTAGTACCAGTTCTCTCTATGGCTAA 1073
 |||||
 Db 188 nAlaAlaThrPheGlyGly-----AlaIleHis 198
 |||||
 QY 1074 CGAGGCTCTTATTCCTTTGTAGCGAATTTGACAGGAGTAGAGGGGGAGGAGTTGCTGC 1133
 |||||
 Db 198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg-----212

QY 1134 TGTTCAGGATGGGCGAGGAGGTGTATCATCTTCAACAGAGATCCAGTAGTAAG 1193
 |||||
 Db 212 -----212
 |||||
 QY 1194 TTTTCCAGAAATACTCGGTAGATTGTGAGGAACTAGCCCGAGTAGGAGGAGGAT 1253
 |||||
 Db 213 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLeu 226
 |||||
 QY 1254 TTACTCTCTACCGGAACTGTCTTCTCTCAATAATAATGGAACAACTTTTCTCAACAATGT 1313
 |||||
 Db 226 uTy-SerAspGlyAspIleAspGlnAsnAlaTyValLeuPheArgGluAsnGl 246
 |||||
 QY 1314 TGCTTCTCTCTTACATTGTCTGTAAAGCAACCAAGTAGGACAGGCTTCTAATACGAG 1373
 |||||
 Db 246 uAlaLeuThrThrAlaIle-----252
 |||||
 QY 1374 TAATAATTACGAGATGCGAGGAGCTATCTTCTGTAAAGATGTTGCGCAAGCAGGATCCAA 1433
 |||||
 Db 253 -----GlyLysGlyAlaValCysCys-----LeuPr 262
 |||||
 QY 1434 TAACCTCTGATCA-----GTTTCTTTGATGAGGAGGAGTATTTT 1475
 |||||
 Db 262 oThrSerGlySerSerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
 |||||
 QY 1476 CTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTATGTCACAAAAGCTCTCGGT 1535
 |||||
 Db 282 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyAlaArgLysLeuSerI 302
 |||||
 QY 1536 TGCTAACTGTGGCTCTGTACAAATTTTAAAGNATATCGCT-----AATCA 1580
 |||||
 Db 302 sSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyAlaAsnSerGlnAsnLe 322
 |||||
 QY 1581 TGGTCGAGCGATTTATTTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTTATGGAGA 1640
 |||||
 Db 322 uGlyGlyAlaIleAlaIleAspThrGlyGlyGluIleSerLeuSerAlaGluLysGlyTh 342
 |||||
 QY 1641 TATTATTTTCGATGGGAATCTTTAAAGAACACGACCAAGAAAGATGTCGCGATGTTAATGG 1700
 |||||
 Db 342 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGl 357
 |||||
 QY 1701 CATAACTGTGCTCTCACAGCCATTTGATGGGATCGGGAGGAGAAATACGACATTAAG 1760
 |||||
 Db 357 yIleHisLeuLeuGlnAsnAla-----LysPheLeuLysLeuGl 370
 |||||
 QY 1761 AGCTAAAGCAGGCGCATCATGTTCTTTAATGATCCATCGAGATGCGCAACGGAATAA 1820
 |||||
 Db 370 nAlaArgAsnGlyTySerIleGluPheTyAspProIle-----Th 384
 |||||
 QY 1821 CCAGCCAGCGCAGTCTTCCAAACTTCTAAAAATTAAC-----GATGGTGAAGATA 1871
 |||||
 Db 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
 |||||
 QY 1872 CACAGGGGATATGTTTTTGTCT-----AATGGAG 1901
 |||||
 Db 404 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheTy 424
 |||||
 QY 1902 CAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGGAGTGTCTTCTGTAAGAGGC 1961
 |||||
 Db 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyLeuValIleLysGluGlyAl 444
 |||||
 QY 1962 AAAATTTATCAGTGAATTTCTTAAGTCAGACAGCTGGGAGT---CTGTATATGGAAGCTGG 2018
 |||||
 Db 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGl 464
 |||||
 QY 2019 GAGTACATGGGATTTTGTAACTCCACACACCAACGCTCTCTGCGCTCAATCAGTT 2078
 |||||
 Db 464 yThr-----LysLeuIleAlaSerLysGluAs 473
 |||||
 QY 2079 GATCAGCTTTCCAATCTGCAATTTGTCTCTTTCTTTCTTTGTTAGCAACAATCCAGTTAC 2138
 |||||
 Db 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu-----486
 |||||
 QY 2139 GAATCCTCTTACCAATCTCCAGCGCAAGATTCTCATCTCTGCAAGTATTGTTAGTACCAAC 2198

487	Db	-----::: :::	SerSerSerThrAlaAlaValIleIysAlaSer	499
2199	QY	-----TCTGTTCAAAATTAGTGGCCCTATCTTT-----	2232	
499	Db	:::	AlaAsnIysGlnIleSerValThrAspSerIleGluIleSerProThrGlyAsnAl	519
2233	QY	-----TTTAGGATTTGGATCATACAGCTTATGATAGGTATGATTGGCTAGGTTCTTAATCAAAA	2291	
519	Db	::: :::	MetArgAsnSerGlnThr	530
2292	QY	-----AATCAATGTCTCTGMAATACATAGTTAGGAGACTAAGCCCCAGCTTAATGCCCCATCAGATT	2351	
530	Db	:::	ProGlyAlaGlyIleSerValThrVal	546
2352	QY	-----GACTTAGGGAATGAGATG-----	2402	
546	Db	:::	IleThrAlaGlyAspPheLeuProValSerProHisTyrGlyPheGlnIleAsnTrpIysLe	566
2403	QY	-----TGGCTGGGATCCTAATACAGCAATAATAGTGTCTTATATACTCTGAAGACTACATGGACTAA	2462	
566	Db	:::	ValLeuTrpThrGlyThrGlyAsnIysValGlyGluPhe-----	582
2463	QY	-----AATCGGGTATAATCTCGGGCTGAGCAGTAGCTTCTTTGGTCCAAATAGTTATAGGGG	2522	
582	Db	:::	SerIleAsnTyrIysProArgProGluIysGluGlyAsnLeuValProAsnIleLeuTrpGly	602
2523	QY	-----ATCCATTTTATAGATATACGATCTGGCATTTCAGCAATTCAGCAAGTGTGGATGGCGCTC	2582	
602	Db	:::	YasnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGly	622
2583	QY	-----TTATTGTCGAGGATTAGGGTTTCTGGAGTTTCGAATTTCTCTCATCATCACCOCGATGC	2642	
622	Db	:::	nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerGly	642
2643	QY	-----TTTAGTTCAGGATATCGGTATATTATGGGGTTTATTCCTTAGGACCAACTCCCTACTT	2702	
642	Db	:::	uAspAsnIleAerGlyrArgHisAsnSerGlyGlyTyrValLeuSerValAsnAsnGluIle	662
2703	QY	-----TGGATCATCGATGTTT---GGTCTAGCATTTACCGAGTATTGGTAGACTAAAGATTA	2759	
662	Db	:::	eThrProIysHisTyrThrSerMetAlaPheSerGlnLeuSerArgAspIysAspTyr	682
2760	QY	-----TGATGTGTGCTTCCCAATCATCATGCTTGCAATGAGTATCGTTTATCTATCTACCCAAACA	2819	
682	Db	:::	AlaValSerAsnAsnGluTyrArg-----MetTyrLeu-----	693
2820	QY	-----AGCTTTATGGATTCCTATTGTTTC-----GGAGATCGGTTTATCCG	2861	
694	Db	:::	GlySerTyrLeuTyrGlnTyrThrThrSerLeuGlyAsnIlePheArgTyr	710
2862	QY	-----TGCTAGC-----	2868	
710	Db		AlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr	730
2869	QY	-----TACGGGTTTGGGAATCAGCATATGAAACCTC	2900	
730	Db		LeuMetIlePheHisPheLeuCysAlaTyrGlyHisAlaThrAsnAspMetIysThrAs	750
2901	QY	-----ATATACATTTGCAGAGGAGCGGATGTTCTGGTGGGATAATACTGTCTGGCTGGAGAGAT	2960	
750	Db		pTyrAlaAsnPheProMetValIysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCys	770
2961	QY	-----TGGAGCGGGATACCGATGTTGATTACTCCATCTAAGCTCTATTGTGAATGAGTTGGCTCC	3020	
770	Db	:::	SerGlyIysMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr	790
3021	QY	-----TTTCGTCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGAACGCCATCA	3080	
790	Db	:::	OpheMetIysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheIysGluThrThrAlaAs	810
3081	QY	-----AGCTCGGGCATTTCAAGCGCGACATCTCCCTAAATCTATCATAGTTCCTGCTTGGATGAGTT	3140	

Db	810	pGlyArgpGpSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPhe	830
Qy	3141	TGATCGATGTTCTAGTACACATCCTTAATAATATAGCTTTATGCGGCTTATATCTGTGA	3200
Db	830	eGlyuLeuAlaLeuSerGlnAspValLeuTyraAspPheSerPheSerTyriLeProAs	850
Qy	3201	TGCTTATCGCACCATCTCTGGTACTGACGACAAACGCTCTCTATCCATCCACGACATGGAC	3260
Db	850	pIlePheArgLyAspProSerCysGluAlaAlaLeuValIleSerGlyAspSerTrpLe	870
Qy	3261	AACAGATGCTTCATTAGACAAGACATGAGTTGTGGTAGAGATCTATGATGCTTC	3320
Db	870	uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyriHi	890
Qy	3321	TCTAACCAAGTAATATAGAAGTATATGCCATGACGAAGTATATGATATCGATGATCTCTCG	3380
Db	890	sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr	910
Qy	3381	AGGCTATGGTTTGAGTGCAGAAAGTAGTCCGGTTC	3417
Db	910	gAsnTyraAsnIleAsnCysGlySerLyAspPheArgPhe	922
RESULT 32			
ID	Q822Q5	CHLCV	
AC	Q822Q5	CHLCV PRELIMINARY;	PRT; 841 AA.
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Polymorphic outer membrane protein G family protein/autotransporter.		
DE	OrderedLocustNames=CCA00624;		
OS	Chlamydomophila caviae.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=83557;		
EN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=GPIC;		
RX	MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;		
RA	Read T.D., Myers G.S.A., Brumham R.C., Nelson W.C., Paulsen I.T.,		
RA	Heidelberg J.F., Holtzaple E.K., Khouri H.M., Federova N.B.,		
RA	Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,		
RA	White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,		
RA	Bavoi P.M., Fraser C.M.;		
RA	"Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):		
RT	examining the role of niche-specific genes in the evolution of the		
RT	Chlamydiaeae."		
EL	Nucleic Acids Res. 31:2134-2147(2003).		
DR	EMBL; A016996; AAP05366.1; -, Genomic_DNA.		
DR	TIGR; CCA00624; -.		
DR	GO; GO:0019867; C:outer membrane; IEA.		
DR	InterPro; IPR006315; Autotransporter.		
DR	InterPro; IPR005546; Auto transportbeta.		
DR	InterPro; IPR011427; ChlamPMP_M.		
DR	InterPro; IPR003368; Chlamydia_PMP.		
DR	Pfam; PF03797; Autotransporter; 1.		
DR	Pfam; PF07548; ChlamPMP M; 1.		
DR	Pfam; PF02415; Chlam_PMP; 4.		
DR	TIGRFAMs; TIGR01414; autotrans_bar1; 1.		
DR	TIGRFAMs; TIGR01376; POMP_repeat; 3.		
XM	Complete proteome.		
SQ	SEQUENCE 841 AA; 90411 MW; 21A9ED16E0D7C65D CRC64;		

Db 495 LeuAlaThrLeuAla---TrpGlnGlnThrGlyTyrAsnProAsnProGluArgGlnGly 513
 Qy 2497 TCTTTGGTTCCTCAATAGTTATGGGATCCATTTTAGATATACGATCTCGGATTCACGA 2556
 Db 514 ProLeuValProAsnThrLeuTrpGlySerPheSerAspValArgAlaIleGlnAsnLeu 533
 Qy 2557 ATTCAAGCAAGTGTGGATGGCGCTCTTATGTCGAGGATATGCGTTCTTGGAGTTTCG 2616
 Db 534 MetAspIleSerValAsnGlyAlaAspTyrGlnArgGlyLeuTrpAlaSerGlyLeuAla 553
 Qy 2617 AATTTCTTCTATCATGACCGCGATCTTTAGCTCAGGATATCGGTATATAGTGGGGT 2676
 Db 554 AsnPheLeuGlnLysSerGlyThrGluThrLysArgLysPheArgHisSerAlaGly 573
 Qy 2677 TATTCCTTAGGACCAATCC---TACTTTGATCATGATGTTGTCTAGCATTTACC 2733
 Db 574 TyrValLeuGlyAlaTyrAlaLysThrLeuSerAspValPheSerAlaAlaPheCys 593
 Qy 2734 GAAGTATTGTGTAGATCTAAGATTTATGATGTCGCTTCCATCATCATGCTTGCGATA 2793
 Db 594 GlnLeuPheGlyArgAspLysTyrLeuValSerLysAsnAsnSerAsnIleTyrAla 613
 Qy 2794 GCATCCGTTTATCTATCTACCAACAAGCTTTATGTCGATCCTATTGTCGAGATCGC 2853
 Db 614 GlySerIleTyrTyrGlnHisThrSer-----PheTrpAspAla 626
 Qy 2854 TTT-----ATCCGT 2862
 Db 627 TrpAspAsnLeuLeuGlnSerThrLeuGlyAlaGlnAlaProLeuValLeuAsnAlaGln 646
 Qy 2863 CTAAGTACGGTGGTGGGATCAGCATATGAACCTCATATACA-----TTTCAGAG 2916
 Db 647 LeuThrTyrSerHisThrSerAsnAspMetLysThrAsnMetThrLysTyrAlaPro 666
 Qy 2917 GAGACCGATGTT-----COTTCGGATATAAATCTGCTCGCTGGAGAG 2958
 Db 667 GlnGlyValValTyrProGluIleLysGlyAspTrpGlyAsnAspCysPheGlyValGlu 686
 Qy 2959 ATTGAGCGGGATACCAATGTTGATTTACTCCATAGCTCTATTGAATGAGTTGGGT 3018
 Db 687 LeuGlyAlaThrValProIleGluSerProTyrSerSerLeuPhe---AspMetTyrSer 705
 Qy 3019 CCTTTCTGTCGACGTTCTTATGCGCATCATGATCTTTTACAGAGGAGCGCAT 3078
 Db 706 ProPheLeuArgPheGlnLeuValTyrAlaHisGlnGluAspPheLysGluAsnAsnSer 725
 Qy 3079 ---CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTCTGGAGTG 3135
 Db 726 ThrGluGlyArgTyrPheGluSerSerAspLeuThrAsnLeuSerMetProIleGlyVal 745
 Qy 3136 AAGTTTGATCGATGTTCTAGTACACATCCTAATAATATAGCTTTATGGCGCTTATATC 3195
 Db 746 LysPheGluArgPheSerAspAsnAspIleAlaSerTyrAsnValThrLeuAlaTyrAla 765
 Qy 3196 TGTGATGTTATCGACCATCTCTGGTACTGAGACAGCTCTCTATCCATCCATCAGAGACA 3255
 Db 766 ProAspLeuValArgSerAsnProAspCysLysThrSerLeuLeuValSerProThrThr 785
 Qy 3256 -----TCGACCAACAGATGCTCTTTCATTTAGCAAGCATGCGATGTTGTGGTAGGATCT 3309
 Db 786 AlaValTrpLeuThrLysAlaThrAsnLeuAlaArgHisAlaPheIleValLysAlaGly 805
 Qy 3310 ATGTATGCTTCTTAACAGTAATATAGAGTATATAGGCCATGGAAGATATGAGTATCGA 3369
 Db 806 AsnTyrLeuSerLeuSerSerAsnPheGluIlePheSerGlnPheGlyPheGluLeuArg 825
 Qy 3370 GATGCTTCTCGAGGCTATGTTGATGTCGAGAGTAGAGTCGGGTC 3417
 Db 826 GlySerSerArgThrTyrAsnValAspLeuGlySerLysIleGlnPhe 841
 RESULT 33
 P71132_CHLAB
 ID P71132_CHLAB PRELIMINARY; PRT; 847 AA.

AC P71132; OS:LSL6J4;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 13-FEB-2005 (TrEMBLrel. 31, Last annotation update)
 DE POMP91A (Polymorphic outer membrane protein).
 GN Name=pmp13G; Synonyms=pomp91A; OrderedLocusNames=CAB281;
 OS Chlamydomonas abortus.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OX NCBI_TaxID=83555;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=S26/3;
 RC MEDLINE=98187897; PubMed=9529048;
 RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
 RT "Molecular cloning and characterization of the genes coding for the
 RT highly immunogenic cluster of 90-kilodalton envelope proteins from the
 RT Chlamydia psittaci subtype that causes abortion in sheep.";
 RL Infect. Immun. 66:1317-1324(1998).
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=S26/3;
 RC MEDLINE=96406378; PubMed=8810511; DOI=10.1016/0378-1097(96)00281-9;
 RA Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
 RT "Identification of a multigene family coding for the 90 kDa proteins
 RT of the ovine abortion subtype of Chlamydia psittaci.";
 RL FEBS Microbiol. Lett. 142:277-281(1996).
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=S26/3;
 RC PubMed=15837807; DOI=10.1101/gr.3684805;
 RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
 RA Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,
 RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
 RA Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;
 RT "The Chlamydomonas abortus genome sequence reveals an array of
 RT variable proteins that contribute to interspecies variation.";
 RL Genome Res. 15:629-640(2005).
 DR EMBL; U65942; AAC15921.1; -; Genomic DNA.
 DR EMBL; CR848038; CAH63731.1; -; Genomic DNA.
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR005546; Auto transportbeta.
 DR InterPro; IPR006315; Auto transporter.
 DR InterPro; IPR011427; ChlamPMP_M.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02415; ChlamPMP; 3.
 DR Pfam; PF07548; ChlamPMP_M; 1.
 DR TIGRFAMs; TIGR01414; autotrans barl; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
 KW Complete proteome.
 SQ SEQUENCE 847 AA; 90695 MW; 754C958E7F11798 CRC64;
 Alignment Scores:
 Pred. No.: 2,77e-54 Length: 847
 Score: 950.00 Matches: 282
 Percent Similarity: 41.5% Conservative: 145
 Best Local Similarity: 27.4% Mismatches: 332
 Query Match: 12.1% Indels: 270
 DB: 2 Gaps: 33
 US-10-701-844-1 (1-4435) x P71132_CHLAB (1-847)
 Qy 490 TACAGTGGGAGACGTTAACTGTATCATTTCCCTATATCTTTATAGGAGATCCGAGTGG 549
 Db 36 TyrAsnGlyAsnThrAsnSerGluProPheAsnProLeuSerThrSerAsnSerAsnGly 55
 Qy 550 ACTACTGTTTCTCGCAGGAGGAGTTAACATTA-----AAAATCTTGCATTTCTATT 603
 Db 56 ThrIleTyrThrCysThrGlyAsnIleCysIleAlaTyrAlaGlyLeuAspGlySer--- 74
 Qy 604 GCAGCTTTCCTTAACTGTTTGGGAACCTATTAGGAGAGTTTACTGTTTATAGGAGAGA 663
 Db 75 --GlyLeuSerSerSerCysPheThrAspThrAlaGlyAsnLeuSerPheLeuGlyAsn 93


```
Db 601 GlnLeuPheGlyAspGlyAspTyrLeuValSerIysAsnSerSerThrValTyrAla 620
Qy 2794 CGATCGGTTTATCTA-----TCTACCCCAACAGCTTTATGGATCC 2835
Db 621 GlySerIleTyrTyrGlnHisIleSerTyrTrpAsnThrTyrPheAsnThrLeuLeuGlnAsn 640
Qy 2836 TATTGTTCCGAGATGCGTTTATC-----CGTGTAGCTACGGGTTGGGAAAT 2883
Db 641 ThrLeuGlyAlaGluAlaProLeuValLeuAsnAlaGlnLeuAlaTyrCysHisAlaSer 660
Qy 2884 CAGCATATAAAACCTCATAT-----ACATTTCGACAGAGGAG-----AGCGAT 2925
Db 661 AsnAsnMetIleThrAsnMetThrAspThrTyrAlaProProLysThrThrTyrSerGlu 680
Qy 2926 GTTCGT-----TGGGATAATAACTGCTCGCTGGAGATTTGGAGCGGGATTACCGATT 2979
Db 681 IleLysGlyAspTyrPheAsnAspCysPheGlyValGluPheGlyAlaLysAlaPro--- 699
Qy 2980 GTGATTACTCATCTAAGCTCTATTGTAATGAGTTGCGTCTTTTCGTCGAAGCTGAGTTT 3039
Db 700 ---IleGluThrAlaSerLeuPheAspMetTyrSerProPheValLysLeuGlnLeu 718
Qy 3040 TCTATGTCGATCATGATCTTTACAG--GAAGCGCATCAAGCTCGGCGATTCAAG 3096
Db 719 ValHisAlaHisGlnAspPheLysGluAsnAsnSerAspGlnGlyArgTyrPheGlu 738
Qy 3097 AGCGGACATCTCTAAATCTATCAAGTCTCTGTTGGAGTGAAGTTTGATPCGATGTTCTAGT 3156
Db 739 SerAsnAsnLeuThrAsnLeuSerMetProIleGlyValLysLeuGluLysPheSerHis 758
Qy 3157 ACACATCTTAATAATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGCACCATC 3216
Db 759 LysAspThrAlaSerTyrAsnLeuThrLeuAlaTyrAlaProAspIleValArgSerAsn 778
Qy 3217 TCTGTCATGAGACACCTCCTA-----TCCCATCAAGACATGACACACAGATGCC 3270
Db 779 ProAspCysThrAlaSerLeuLeuValSerProThrSerAlaValTrpValThrLysAla 798
Qy 3271 TTTTCATTAGCAAGACATGAGGATGTTGTGTTAGAGGATCTATGATGCTTCTCTAACAGT 3330
Db 799 AsnAsnLeuAlaArgHisAlaPheIleLeuGlnAlaGlyAsnTyrLeuAlaLeuThrArg 818
Qy 3331 AATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCTCGAGGCTTGGT 3390
Db 819 AsnThrGluLeuPheSerGlnPheGlyPheGluLeuArgGlySerCysArgThrTyrAsn 838
Qy 3391 TTGAGTCGACGAAGTAGATCGGTTTC 3417
Db 839 IleAspLeuGlySerLysIleGlnPhe 847

RESULT 34
ID P77792_CHLAB PRELIMINARY; PRT; 839 AA.
AC P77792_QSL595;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 13-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE POMP90B precursor (POMP90A precursor) (Polymorphic outer membrane
DE protein)
GN Name=pmp17G; Synonyms=pomp90B; OrderedLocusNames=CAB598;
OS Chlamydomonas abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomonadophila.
OC NCBI_TaxID=83555;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S26/3;
RX MEDLINE=98187897; PubMed=9529048;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "Molecular cloning and characterization of the genes coding for the
RT highly immunogenic cluster of 90-kilodalton envelope proteins from the
RT Chlamydia psittaci subtype that causes abortion in sheep."
RL Infect. Immun. 66:1317-1324(1998).
```

```
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S26/3;
RX MEDLINE=96406378; PubMed=8810511; DOI=10.1016/0378-1097(96)00281-9;
RA Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RT "Identification of a multi-gene family coding for the 90 kDa proteins
RT of the ovine abortion subtype of Chlamydia psittaci."
RL FEMS Microbiol. Lett. 142:277-281(1996).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S26/3;
RX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingston M., Cerdano-Iarraga A.-M., Harris B., Doggett J.,
RA Quimond D., Mungall K., Clarke K., Fellwell T., Hance Z., Sanders M.,
RA Ormrod M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomonas abortus genome sequence reveals an array of
RT variable proteins that contribute to inter-species variation."
RL Genome Res. 15:629-640(2005).
DR EMBL; U65943; AAC15924.1; -; Genomic_DNA.
DR EMBL; U65942; AAC15922.1; -; Genomic_DNA.
DR EMBL; CR848038; CAH64045.1; -; Genomic_DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transporter.
DR InterPro; IPR003315; Auto transporter.
DR InterPro; IPR011427; Chlamydia_PMP.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam PMP; 4.
DR Pfam; PF07548; ChlamPMP M; 1.
DR TIGRPFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRPFAMs; TIGR01376; POMP_repeat; 3.
DR Complete proteome; Signal.
FT SIGNAL 1 16 Potential.
FT CHAIN 17 839 POMP90B.
SQ SEQUENCE 839 AA; 89825 MW; 4581C7CBAP7FF4C4 CRC64;

Alignment Scores:
Pred. No.: 2,99e-54 Length: 839
Score: 949.50 Matches: 291
Percent Similarity: 41.7% Conservative: 153
Best Local Similarity: 27.3% Mismatches: 341
Query Match: 12.0% Indels: 281
DB: Gaps: 36

US-10-701-844-1 (1-4435) x P77792_CHLAB (1-839)
Qy 382 ATGCAAAAGCTTTTCCATAAGTTCTTTTCAATGATTCTAGCTATTCTTGCTGCTCT 441
Db 1 MetLysHisProValTyrTrpPheLeuIleSerSerSerLeu---PheAlaSerAsnSer 19
Qy 442 TTAATGGGGGGGATATGCAGCAGAAATCATGGTCTCTCAAGGAATTTACGATGGGAG 501
Db 20 LeuSerPheAlaAsnAspAlaGlnThrAlaLeuThrProSerAspSerTyrAsnGlyAsn 39
Qy 502 ACGTTAATGTATTCATTTCCCTATATCTGTTATAGGAGATCCGATGGGACTACTGTTTT 561
Db 40 ValThrSerGluGluPhe-----GlnValLysGluThrSerSerGlyThrThrTyrThr 57
Qy 562 TCTGCAGGAGGTAAACATTAATAAAATCTTGACAATTTCTATTGCGAGCTTTCGCTTAAGT 621
Db 58 CysGluGlyAsnValCysIle---SerPheAlaGlyLysAspSerGlyLeuLysLysSer 76
Qy 622 TGTTTTGGGAACCTTATTAGGAGGTTTACTGTTTTAGGGAGAGACACTCGTTGACTTTC 681
Db 77 CysPhe---SerAlaThrAspAsnLeuThrPheLeuGlyAsnGlyTyrThrLeuCysPhe 95
Qy 682 GAGAACATACGACTTCTACAAATGGGGGAGCTCTAAGTAAATGCGCTGCTGATGGA--- 738
Db 96 AspAsnIleThrThrThrAlaSerAsnProGlyAlaIleAsnValGlnGlnGlyLys 115
Qy 739 CTGTTTACTATTGAGGGGTTTAAAGATTAATCTCTTTTCCCAATTTCATTACTTGGC 798
```


Db 116 ThrLeuGlyLeuSerGlyPheSerLeuPheSerCysAlaTyrCys----- 130
 QY 799 GTACTGCTGCTGCAACGACTAATAAGGTAGC---CAGACTCCGACGACCAACCTACA 855
 Db 131 -----ProProGlyThrThrGlyTyrGlyAlaIleGlnThrLysGlyAsnThrLeu 148
 QY 856 CCGTCTAATGGTACTATTATTCTTAACAAACAGATCTTTTGTACTCAATAAGTGAAGTTC 915
 Db 149 LysAspAsnSerSerLeu----- 154
 QY 916 TCATTCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACG 975
 Db 154 ----- 154
 QY 976 GTTCAAGGAATTAGCAAGCTTTGTCTTCCAGAAAATACTGCTCAAGCTGATGGGCA 1035
 Db 155 -----ValPheHisLeuAsnCysSerThrAlaGluGlyGly 166
 QY 1036 GCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCTTTGTA 1095
 Db 167 AlaIleGln----- 169
 QY 1096 GCGAATGTTGCAGAGTAGTAAGAGGGGAGGAGTTGCTGCTTTCAGGATGGCGACAGGGA 1155
 Db 169 ----- 169
 QY 1156 GTGTCATCATCTACTTCAACAGAGAATCCAGTAGTAAGTTTTCAGAAATACTGCGGTA 1215
 Db 169 ----- 169
 QY 1216 GAGTTTGTGAGGAACTGAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCT 1275
 Db 169 ----- 169
 QY 1276 TTCCTGAATAATGMAAAACCTTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTGCT 1335
 Db 169 ----- 169
 QY 1336 GCTAAGCAACCAAGTGGACAGGCTTCTAATACGAGTAATAATACGAGATGGAGGA 1395
 Db 169 ----- 169
 QY 1396 GCTATCTCTGTAAGAATGTGCGCAGCAGGA-----TCCATAACTCTGGATCA 1446
 Db 170 -----CysLysGlySerSerAspAlaGluLeuLysIleGluAsnAsnGlnAsnLeu 186
 QY 1447 GTTTCCTTTGATGGAGGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGGAAAGGG 1506
 Db 187 Val-----PheSerGluAsnSerSerThrSerLysGly 197
 QY 1507 GGAGCTATTTATGCCAAAAGCTCTCGTTGCTTAAGTGTGCTGCTGACAAATTTTAAAGG 1566
 Db 198 GlyAlaIleTyrAlaAspLysLeuThrIleValSerGlyGlyProThrLeuPheSerAsn 217
 QY 1567 AATATCGCTAATGAT-----GGTGGAGCGATTTATTATTAGGAA--TCT 1608
 Db 218 AsnSerValSerAsnGlySerSerProLysGlyGlyAlaIleSerIleLysAspSerSer 237
 QY 1609 GGAGAGCTCAGTTTATCTGCTGATTATGAGATATTATTTTCGATGGGAAT---CTTAAA 1665
 Db 238 GlyGluCysSerLeuThrAlaAspLeuGlyAspIleThrPheAspGlyAsnLysIleIle 257
 QY 1666 AGAACAGCAAAAGAGATGCTGCGGATGTTAATGGCTAACTGTGCTGCTCAACAGCCATT 1725
 Db 258 LysThrSerGlyLysSer-----ThrValThrArgAsnSerIle 271
 QY 1726 TCGATGGATCGGAGGGAAAATAACGACATTAAAGAGTAAAGCAGGGCATCAGATTCTC 1785
 Db 272 AspLeuGlyThr--GlyLysPheThrLysLeuArgAlaLysAspGlyPheGlyIlePhe 290
 QY 1786 TTTAATGATCCATCGAGATGGCAACCGAATAACGAGCGGCGAGCTCTTCCAAAACCTT 1845
 Db 291 PheTyrAspProIle--ThrGlyGlySerAspGlu----- 302

QY 1846 CTAAAAATTAACGATGGTGA-----GGATACACAGGGGATATTGTTTTTCT----- 1893
 Db 303 LeuAsnIleAsnLysLysGluThrValAspTyrThrGlyLysIleValPheSerGlyGlu 322
 QY 1894 -----AATGAAGCAGTACTTTGTACCAAAAT 1920
 Db 323 LysLeuSerAspGluLysAlaArgAlaGluAsnLeuAlaSerThrPheAsnGlnPro 342
 QY 1921 GTTACGATAGACAGGAAGGATTTCTTCTGTAAGAAAGGCAAAATTTACGTGAATTC 1980
 Db 343 IleThrLeuSerAlaGlySerLeuValLeuLysAspGlyValSerValThrAlaLysGln 362
 QY 1981 CTAACTCAGACAGGGGAGT---CTGTATATGGAAGCTGGAGTACATGGGATTTGTA 2037
 Db 363 ValThrGlnGluAlaGlySerThrValValMetAspLeuGlyThrThrLeu----- 379
 QY 2038 ACTCACAAACACCAACAGCCTCTCGCGCTAATCAGTTGATCAGCGCTTTCCATCTCG 2097
 Db 380 -----GlnThrProSerSerGlyGlyGluThrIleThrLeuThrAsnLeu 394
 QY 2098 CATTTGCTCTTTCTTTTGTAGCAAAATGCAATGCAATGCAATGCTCTCTACCAATCCT 2157
 Db 395 AspIleAsnIleAlaSerLeuGlyGlyGlyGlyThrSer----- 408
 QY 2158 CCAGCGCAAGATCTCTCATCTCGCAGTCAATTGGTAGCACAACTGCTGGTTCTGTACAAT 2217
 Db 409 -----ProAlaLysLeuAlaThrAsnThrAlaSerGlnAlaIleThr 422
 QY 2218 AGTGGCGCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGGATGATGGCTA 2277
 Db 423 IleAsnAlaValAsnLeuValAspAlaAspGlyAsnAlaTyrGlu-----AspProIle 440
 QY 2278 GTTCTTAATCAAAATCAATGCTCTGAAATACAGTTAGGAGTAAAGCCCCAGCTAAT 2337
 Db 441 LeuAlaThrSerLysProPheThrAlaIleValAlaThrThrAsnAlaSerThrValThr 460
 QY 2338 GCCCATCAGATTGACTCTAGGAAATGAGATGCCT-----AAGTATGCTATCAAGGA 2391
 Db 461 GlnProThrAsp---AsnLeuThrAsnTyrValProThrHisTyrGlyThrGlnGly 479
 QY 2392 AGCTGGAAGCTTCGCTGGGATCCTAATACAGCAAAATAATGGTCTTATCTCTGAAGCT 2451
 Db 480 AsnTyrThrValThrTrpAspThrGluThrAlaThrLys-----ThrAlaThrLeu 496
 QY 2452 ACATGCACTAAACTGGTATATCTCGGCTGAGCAGTAGCTCTTTTGGTTCCAAAT 2511
 Db 497 ThrTrpGluGlnThrGlyTyrSerProAsnProGluArgGlnGlyProLeuValProAsn 516
 QY 2512 AGTTTATGGGATCCATTTTAGATATACGATCTGCGCATTCAGCAATTCAGCAAGTGTG 2571
 Db 517 ThrLeuTrpGlyAlaPheSerAspLeuArgAlaIleGlnAsnLeuMetAspIleSerVal 536
 QY 2572 GATGGCGCTCTTTATTGTCAGGATTAATGGTTTCTGGAGTTTCGAATTTCTTCTATCAT 2631
 Db 537 AsnGlyAlaAspTyrHisArgGlyPheTrpValSerGlyLeuAlaAsnPheLeuHisLys 556
 QY 2632 GACCGGATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTATTCCTTAGGAGCA 2691
 Db 557 SerGlySerAspThrLysArgLysPheArgHisAsnSerAlaGlyTyrAlaLeuGlyVal 576
 QY 2692 AACTCTCTCTTTGGATCA---TCGATGTTTGGTCTAGCATTTACCGAAGTATTGGTGA 2748
 Db 577 TyrAlaLysThrProSerAspAspIlePheSerAlaAlaPheCysGlnLeuPheGlyLys 596
 QY 2749 TCTAAAGATATGATGATGCTGCTGTTCCAAATCATCATCTGTCATAGGATCCGTTATCTA 2808
 Db 597 AspLysAspTyrLeuValSerLysAsnAsnAlaAsnIleTyrAlaGlySerLeuTyr 616
 QY 2809 -----TCTACCCAAACAGCTTTATGTGATCTTATTTGTCGGAGAT 2850
 Db 617 GlnHisIleSerTyrTrpSerAlaTrpGlnAsnLeuGlnAsnThrIleGlyAlaGlu 636

```

QY 2851 GGGTTATC-----CGTGCTAGTACGGGTTTGGGATCAGCATCATGAAACC 2898
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 637 AlaProLeuValLeuAsnAlaGlnLeuThrTyCysHisAlaSerAsnAspMetLysThr 656
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 2899 TCATAT-----ACATTTCGACAG---GAGAGCGCATGTT 2928
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 657 AsnMetThrThrThrTyAlaProArgLysThrThrTyAlaGluLeuLysGlyAsp--- 675
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 2929 CGTTGGGATATAACTGCTGCTGGCGAGATGGAGCGGATACCGAATTGTGATTACT 2988
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 676 ---TTPGlyAsnAspCysPheGlyValGluLeuGlyAlaThrValProIleGln---Thr 693
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 2989 CCATCAGCTCTATTGGAATGAGTTCGTCCTTCGTCGCAAGCTGAGTTCCTTATGCC 3048
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 694 GluSerSerLeuLeuPheAspMetTyrSerProPheLeuLysPheGlnLeuValHisThr 713
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3049 GATCATGATCTTTTACAGC---GAAGCGCATCAAGCTCGGGCATTCAGAGCGGCAT 3105
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 714 HisGlnAspAspPheLysGluAsnAsnSerAspGlnGlyArgTyrPheGluSerSerAsn 733
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3106 CTCCTAAATATCATAGTTCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCT 3165
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 734 LeuThrAsnLeuSerLeuProIleGlyIleLysPheGluArgPheAlaAsnAsnAspThr 753
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3166 AATAAATATAGCTTTATCGGGCTTATCTGTGATGCTTATCGCACCATCTCTGGTACT 3225
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 754 AlaSerTyHisValThrAlaAlaTySerProAspIleValArgSerAsnProAspCys 773
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3226 GAGACAAAGCTCTCTATCCATCAAGACACA-----TGGACACAGATGCGCTTCATTTA 3279
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 774 ThrThrSerLeuLeuValSerProAspSerAlaValTrpValThrLysAlaAsnAsnLeu 793
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3280 CAAGACATGAGTGTGTTAGAGGATCTATGATGCTTCTTAAACAGTAATATAGAA 3339
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 794 AlaArgSerAlaPheMetLeuGlnAlaGlyAsnTyLeuSerLeuSerHisAsnIleGlu 813
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3340 GTATATGCCATGGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGTGATCA 3399
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 814 IlePheSerGlnPheGlyPheGluLeuArgGlySerSerArgThrTyAsnValAspLeu 833
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 3400 GGAAGTAGATCCGGTTC 3417
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 834 GlySerLysIleGlnPhe 839
Db ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 35
PMP2 CHLPN
ID PMP2 CHLPN STANDARD; PRT; 841 AA.
AC Q923A1; Q9RB73;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp2 precursor (Polymorphic membrane
DE protein 2) (Outer membrane protein 7).
GN Name=pmp2; Synonyms=omp7;
GN OrderedAccession=CP0013; CP0761; CPB0015/CPB0016;
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83558;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity";
RL Am. Heart J. 138:8491-8495(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;

```

```

RT *Comparative genomes of Chlamydia pneumoniae and C. trachomatis.;
RL Nat. Genet. 21:385-389(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bas S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20030349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RX Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -1- CAUTION: Ref.4 and Ref.5 sequences differ from that shown due to a
CC frameshift in position 673.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ133035; CAB37083.1; -; Genomic DNA.
CC EMBL; AE001586; AAD18172.1; -; Genomic DNA.
CC EMBL; AE002235; AAF38561.1; -; Genomic DNA.
CC EMBL; BA000008; BAA98223.1; ALT_FRAME; Genomic DNA.
CC EMBL; AE017157; AAP97948.1; ALT_FRAME; Genomic DNA.
CC EMBL; AE017157; AAP97949.1; ALT_FRAME; Genomic DNA.
CC PIR; E72130; E72130.
CC PIR; CP0761; -.
CC InterPro; IPR005546; Auto_transp_beta.
CC InterPro; IPR011427; ChlamPMP_M.
CC InterPro; IPR003368; Chlamydia_PMP.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; Chlam_PMP; 4.
CC Pfam; PF07548; ChlamPMP_M; 1.
CC TIGRFAMs; TIGR01376; POMP_repeat; 3.
CC Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
CC SIGNAL
CC CHAIN 25 841 Probable outer membrane protein pmp2.
CC CONFLICT 784 784 N -> S (in Ref. 4).
CC SEQUENCE 841 AA; 89601 MW; 9064D60D0678D24C CRC64;
CC -----
CC Alignment Scores:
CC Pred. No.: 1.38e-53 Length: 841
CC Score: 939.50 Matches: 289
CC Percent Similarity: 39.5% Conservative: 134
CC Best Local Similarity: 27.0% Mismatches: 358
CC Query Match: 11.9% Indels: 291
CC DB: 1 Gaps: 30

```


Db 127 SerCysAlaTyxCys-----CysProProGlyThrThrGlyTyxGly 140
Qy 829 AGC---CAGACTCCGACGACAACTCTACACGCTCTAATGGTACTATTATTCTAAACA 885
Db 141 AlalleGlnThrLysGlyThrThrThrLysAspAsnSerLeu----- 156
Qy 886 GATCTTTGTTACTCAATAATCAGAAAGTCTCTATTCTATAGTAATTTAGTCTCTGGAGAT 945
Db 156 ----- 156
Qy 946 GGGGAGCTATAGATCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTCTTC 1005
Db 157 -----ValPhe 158
Qy 1006 CAAGAAATACGTCTCAAGCTGATGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCT 1065
Db 159 HisLysAsnCysSerThrAlaGluGlyAlaIleGln----- 171
Qy 1066 ATGGCTAACGAGGCTCTATTGCTTTGTAGCGAATGTTTCAGGAGTAAGAGGGGAGGG 1125
Db 171 ----- 171
Qy 1126 ATTGCTGCTTTCAGGATGGCAGCAGGAGTGTCATCTACTTCAACAGAAATCCA 1185
Db 172 -----CysLysSerSerSerThrAlaGlu--- 180
Qy 1186 GTAGTAAGTTTTTCCAGAAATACTGCGGTAGAGTTGTAGGGAACTAGCCCGAGTAGGA 1245
Db 180 ----- 180
Qy 1246 GGGGAGTTACTCTCTACGGGAACGTTGCTTCTGAATAATGAAAAACCTTGTCTC 1305
Db 181 -----LeuLysLeuGluAsnAsnLysAsnLeu----- 189
Qy 1306 AACAAATGTTGCTTCTCTGTTTACATTGCTGCTAAGCAACCAAGTAGCGACAGGCTTCT 1365
Db 189 ----- 189
Qy 1366 AATACGAGTAATAATTACGAGATGCGAGAGCTATCTTCTGTAAGAATGTCGCGAAGCA 1425
Db 189 ----- 189
Qy 1426 GGNATCCAACTCTCGATCAGTTTCTTTGATGGAGGAGTAGTTTCTTTAGTAGC 1485
Db 190 -----ValPheSerGlu 193
Qy 1486 AATGTAGTCTCGGAAAGGGGAGCTATTATTGCCAAAAGCTCTCGGTTGCTAACTGT 1545
Db 194 AsnSerSerLysGluLysGlyAlaIleTyThrAlaAspLysLeuThrIleValSerGly 213
Qy 1546 GGCCTGTACAAATTTTAAAGGAATATCGCTAAT-----GATGGTGAGCG 1590
Db 214 GlyProThrLeuPheSerAsnAsnSerValSerHisAsnSerSerProLysGlyGlyAla 233
Qy 1591 ATTATTATAGGAATCT---GGAGACTCAGTTTATCTGCTGATATGAGATATATT 1647
Db 234 IleCysIleLysAspSerAspGlyGluCysSerLeuThrAlaAsnLeuGlyAspIleThr 253
Qy 1648 TTCATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCGCATGTTAATGGCGTAAC 1707
Db 254 PheAspGlyAsn-----LysIleIleThrThrAsnGlyLysSerProThr 268
Qy 1708 GTGTCTCTCACAAAGCCATTTGATGGGATCGGAGGGAATAAACGACATTAAGAGCTAAA 1767
Db 269 ValThrArgAsnSerIleAspLeuGlySerGlyLysPheThrLysLeuAsnAlaLys 288
Qy 1768 GCAGGCAATCAGATCTCTTAATGATCCCATCGAGATGGCAACGAAATAACCAAGCA 1827
Db 289 GluGlyPheGlyIlePhePheTyxAspProIleAlaAsnThrGlyLysThrGlu--- 307
Qy 1828 GCGAGTCTTCCAACTCTTAAATTAACGATGGTGAAGATACACAGGGGATTTGTT 1887
Db 308 -----IleGluLeuAsnLysThrGluSerAspThrThrThrThrGlyLysIleVal 324

Qy 1888 TTT-----GCTAATGGAGCAGTACT 1908
Db 325 PheSerGlyGluLysLeuSerAspGluGluLysThrValProAlaAsnLeuLysSerTyx 344
Qy 1909 TTGTACCAAAATGTTACGATAGACAGGAAGGATTGTTCTTCGTGAAGGCAAAATTA 1968
Db 345 PheLysGlnProLeuLysIleGlyAlaGlySerLeuValLeuLysAspGlyValThrLeu 364
Qy 1969 TCAGTGAATTTCTTAAGTCACAGAGTGGGAGT---CTGTATATGGAAGTCGGAGTACA 2025
Db 365 GluAlaLysLysIleThrGlnThrLysGlySerThrValValMetAspLeuGlyThrThr 384
Qy 2026 TGGGATTTTGAATCTCCAAACCAACAGCCTCTCGCGCTTAATCAGTTGATCAGG 2085
Db 395 Leu-----GlnThrProSerSerSerGlyGluThrIleThr 396
Qy 2086 CTTTCCAATCTGCATTTGCTCTTCTTCTTTGTTGTTAGCAACAATGCGAGTTACGAATCCT 2145
Db 397 LeuThrAsnLeuAspIleAsnIleAlaSerLeuGlyGlyGlyThrAla----- 414
Qy 2146 CCTACCAATCTCCAGCGCAAGATTCTCATCTGCAGTCATTGGTAGCACAACTGCTGCT 2205
Db 415 -----ProAlaLysLeuAlaThrAsnThrAlaSer 424
Qy 2206 TCTGTTACAATTTAGTGGGCTATCTTTTTCAGGATTTGGATGATACAGCTTTATGATAG 2265
Db 425 GlnAlaIleSerIleAlaAlaValAsnLeuValAsnThrAspSerAsnThrTyxGlu--- 443
Qy 2266 TATGATGGCTAGGTTCTAATCAAAAATCAATGCTCGAAATACAGTTAGGACTAG 2325
Db 444 ---AspProIleLeuSerAlaSerLys---SerPheSerAlaIleThrAlaThrSer 461
Qy 2326 CCCCAGCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCT-----AAGTAT 2379
Db 462 SerSerThrValThrProGluThrAsnLeuLysAsnTyxThrProThrHisTyx 481
Qy 2380 GCTATCAAGGAAGCTGGAAGCTTCGTGGGATCTCTAATACACGAAATTAATGTCCTTAT 2439
Db 482 GlyTyxGlnGlyAsnTrpThrValThrTrpLysGlnGlySerSerAlaGln---GluLys 500
Qy 2440 ACTCTGAAGCTACATGGACTAAACTGGGTATATCTCGGCTCGAGGAGTAGTCTTCT 2499
Db 501 ThrAlaThrLeuThrTrpGluGlnThrGlyTyxSerProAsnProGluArgValGlySer 520
Qy 2500 TTGTTCCAAATAGTTTATGGGATCCATTTTATAGATATACGATCTCGCAATCAGCAAT 2559
Db 521 LeuValProAsnThrLeuTrpGlyAlaPheSerAspThrArgAlaIleGlnAsnLeuMet 540
Qy 2560 CAAGCAAGTGTGATGGGCTCTTATTCTCGAGGATTTATGGGTTTCTCGAGTTTCCAAAT 2619
Db 541 AspIleSerValAsnGlyAlaAspTyxSerArgGlyPheTrpValSerSerLeuAlaAsn 560
Qy 2620 TTCTCTCATCATGACCGCATGCTTTAGTTCAGGATATCGGTATATTATGTTGGGTTAT 2679
Db 561 PheLeuAsnLysSerGlySerAspThrLysArgLysPheArgHisSerAlaGlyTyx 580
Qy 2680 TCCTTAGGACAACTCTCTACTTTTGGATCA---TCGATGTTTGGTCTAGCATTTACGAA 2736
Db 581 AlaLeuGlyValTyxAlaGlnThrProSerAspAspValCysSerAlaAlaPheCysGln 600
Qy 2737 GTATTTGGTAGATCTAAAGATTATGTTGTGTTTCAATCATCATCTGTCATAGGA 2796
Db 601 LeuPheGlyLysAspLysAspTyxPheValSerLysAsnSerSerThrIleTyxAlaGly 620
Qy 2797 TCGTTTATCTA-----TCTACCCCAACAGCTTTTATGTGATCTCTAT 2838
Db 621 SerIleTyxTyxGlnHisIleSerTyxTrpAsnThrTrpAsnThrLeuLeuGlnAsnThr 640
Qy 2839 TTCTTCGGAGATCGGTTTATC-----CGTCTAGCTACCGGTTTGGGAATCAG 2886
Db 641 LeuGlyAlaGluAlaProLeuValLeuAsnAlaGlnLeuThrTyxTyxHisAlaSerAsn 660


```

QY 2887 CATATGAAACCTCATATACA-----TTTGCAGAGGAGCGATGTT 2928
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 AsnMetLysThrAsnMetThrAsnThrThrProLysAsnValThrProSerGluIle 680
QY 2929 CGT-----TGGGTAATAACTGCTGCTGCGAGAGATTGGAGCGGGATTACCGATTGTG 2982
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
681 LysGlyAspTTPGlyAsnAspCysPheGlyValGluPheGlyAlaLysAlaPro----- 698
QY 2993 ATTACTCATCTAGCTCTATTGTAATGAGTTGGCTCTTTGTCGCAAGCTGAGTTTCT 3042
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
699 IleGluThrAlaSerLeuPheAspMetTyrSerProPheValLysLeuGlnLeuVal 718
QY 3043 TATGCGGATCATGATCTTTTACAGAG---GAAGCGGATCAAGCTCGGGCATTCAAGAGC 3099
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
719 HisAlaHisGlnAspAspPheLysGluAsnAsnSerAspGlnGlyArgTyrPheGluSer 738
QY 3100 GGACATCTCCCTAAATCTATAGTTCTCTTGGAGTGAAGTTTGGATCGATGTTCTAGTACA 3159
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
739 AsnAsnLeuThrAsnLeuSerMetProIleGlyValLysLeuGluLysPheSerHisLys 758
QY 3160 CATCTTAATAATATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCT 3219
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
759 AspThrAlaSerTyrAsnLeuThrLeuAlaTyrAlaProAspIleValArgSerAsnPro 778
QY 3220 GGTACTGAGACACGCTCTTA-----TCCCATCAAGAGACATGCACACAGATGCCCTTT 3273
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
779 AspCysThrAlaSerLeuLeuValSerProThrSerAlaValTrpValThrLysAlaAsn 798
QY 3274 CATTTAGCAAGACATGGAGTTGTGTAGAGGATCTATGTCTTCTTAACAAGTAAT 3333
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
799 AsnLeuAlaArgHisAlaPheIleLeuGlnAlaGlyAsnTyrLeuAlaLeuThrArgAsn 818
QY 3334 ATAGAAGTATATGGCCATCGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTG 3393
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
819 ThrGluLeuPheSerGlnPheGlyPheGluLeuArgGlySerCysArgThrTyrAsnIle 838
QY 3394 AGTGCAGGAAGTAGATCCGGTTC 3417
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
839 AspLeuGlySerLysIleGlnPhe 846

RESULT 37
Q823X2_CHLVC PRELIMINARY; PRT; 843 AA.
AC Q823X2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family protein/autotransporter,
DE putative.
GN OrderedLocusNames=CCNA00281;
OS Chlamydomydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.P., Holtzapple E.K., Khouri H.M., Federova N.B.,
RA Carty H.A., Unayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Haia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydomydia caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147 (2003).
DR EMBL; AE016995; AAP05032.1; -, Genomic_DNA.
DR TIGR; CCA00281; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transpbteta.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.

```

```

DR Pfam; PF07548; ChlamPMP M; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR TIGRFAMS; TIGR01376; POMP_repeat; 3.
KW Complete proteome.
SQ SEQUENCE 843 AA; 91510 MW; 05CCE3764CF3A43 CRC64;

Alignment Scores:
Pred. No.: 1.61e-44 Length: 843
Score: 803.00 Matches: 270
Percent Similarity: 38.1% Conservative: 137
Best Local Similarity: 25.3% Mismatches: 361
Query Match: 10.2% Indels: 300
DB: 2 Gaps: 36

US-10-701-844-1 (1-4435) x Q823X2_CHLVC (1-843)
QY 418 ATTCTAGCTTATCTTCTGCTCTTTAAATGGGGGGGATATGCGAGCAATCATGTT 477
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ValLeuLeuPheSerSerPheAlaLeuSer-----IleAlaThrGluLeuAla 24
QY 478 CCTCAAGAAATTTACGATGGGAGACGTTAACTGTA----- 513
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 -----AspAlaAspThrValAsnLeuAlaAlaGlyPheAsnGlySerSer 39
QY 514 -----TCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTGTTTCTGCA 567
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 SerGluThrPheAsnValLysGlnThrAsnAsnValGluGlyThrThrTyrThrLeuThr 59
QY 568 GGAGAGTTTAACTTAAATAAATCTTGACAAATCTTATTGCAGCTTTGCCTTTTAAAGTTGTTT 627
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 GlyAlaValSerPheAlaAsnIle---AsnLysPheAspGlnAlaAspThrSerCysPhe 78
QY 628 GGGAACTTATPAGGAGTTTATCTGTTTGGGAGAGGACACTCGTTGACTTTCGAGAAC 687
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 AlaAsnSerAlaGlyAspLeuThrPheThrGlySerArgArgLeuLeuTyrPheAsnAsn 98
QY 688 ATACGGACTTCTACAAATGGGCGAGCTTAAGTAATAGCGCTGCTGATGGACTGTTTACT 747
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
99 IleThrSerSerAlaLysGlyAlaIleSerThrThrAla----- 112
QY 748 ATTGAGGGTTTAAAGAAATTTATCTTTTCCAAATTTGCAATTCATTACTTTCGCTACTGCT 807
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 -----AspAlaLysThrLeuThrLeuSerGlyCysLeuSerLeuIlePheTyrMet--- 129
QY 808 GTCGCAACGACTAATAAGGTAGCCAGACTCCGACGACACATCTACACCGCTCTAATGCT 867
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 -----SerProLysGluAspIleGlyAsnGly 138
QY 868 ACTATTTATTTCTAAAACAGATCTTTTCTTACTCAATATGAGAAGTTCTCATCTATAGT 927
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 AlaIleTyrSerAsnSerMetLeuIleGlnAsnSerAspValSerPheGlyTyr 158
QY 928 AATTTAGTCTCTGAGAGTGGGGAGCTATAGATGCTAAGAGCTTAAAGCGTTTCAAGGAAT 987
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
159 AsnLysSerAlaGlyLysGlySerValIle----- 168
QY 988 AGCAAGCTTTGTGTTCTTCCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTA 1047
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 -----LeuCysGluLys 172
QY 1048 GTACACGATTCTCTGCTATGGCTTAACGAGGCTCTATTGCTTTGTAGCGAATGTTGCA 1107
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 SerThrSer----- 175
QY 1108 GGAGTAAGGGGGGAGGATGCTGCTGTTTTCAGGATGGGACGAGGAGTGTCTCATCT 1167
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 -----AlaGlyAlaThrSerPro 181
QY 1168 ACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTCAGAAATACTCGCGTAGAGTTGATGG 1227
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 ThrLeuThr----- 184
QY 1228 AACGTAGCCCGAGTAGGAGGAGGATTACTCTCTACGGGAACGTTGCTTTCTCTGAATAAT 1287

```


RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LLG;
 RA Laroucau K., Souriau A., Rodolakis A.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF243418; AAL36962.1; -; Genomic_DNA.
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR006315; Auto:transporter.
 DR InterPro; IPR005546; Auto:transporter.
 DR InterPro; IPR011427; ChlamPMP M.
 DR Pfam; PF03797; Auto:transporter; 1.
 DR Pfam; PF07548; ChlamPMP M; 1.
 DR TIGRFAMs; TIGR01414; autoctrans_bar1; 1.
 FT NON_TER
 SQ SEQUENCE 602 AA; 65561 MW; CA486CFACEC131E2 CRC64;

Alignment Scores:

Pred. No.: 1.57e-43 Length: 602
 Scores: 787.50 Matches: 210
 Percent Similarity: 48.6% Conservative: 99
 Best Local Similarity: 33.0% Mismatches: 250
 Query Match: 10.0% Indels: 77
 DB: 2 Gaps: 21

US-10-701-844-1 (1-4435) x Q8VU49_CHLPS (1-602)

QY 1636 GGAGATATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTGCCGATGTT 1695
 DB 2 GlyAspIleThrPheAspGlyAsnLysIleIleThrThrSerArgSerSer----- 19
 QY 1696 AATGCGTAACGTGCTCTCACAAGCCATTTGATGGATCGGAGGAGAAATTAACGACA 1755
 DB -----ThrValLysArgAsnSerIleSerLeuGlySerGlyLysPheThrLys 36
 QY 1756 TTAAGAGCTAAAGCAGGCGCATCTCTTAAATGATCCATCGAGATGCAAAACGGA 1815
 DB 37 LeuAsnAlaLysGluGlyPheGlyIlePhePheThrAspProIleAlaAsnThrGlyAsp 56
 QY 1816 AATAACCGCCAGCGCAGCTCTTCCAAACTCTTAAAAATTAACGATGGTGAAGGA----- 1869
 DB -----ThrAsnThrGlu-----IleGluLeuAsnLysAlaGluGlyLys 70
 QY 1870 -----TACACAGGGATATGTTTTGCT----- 1893
 DB 71 ThrThrThrThrGlyLysIleValPheSerGlyGlyLysLeuSerAspGluGlyLys 90
 QY 1894 -----AATGGAAGCAGTACTTTGACCAAAATGTTACGATAGACCAAGGAGGATT 1944
 DB 91 ValAlaAspAsnLeuLysSerThrPheThrGlnProLeuLysIleGlyAlaGlySerLeu 110
 QY 1945 GTTCTCTGTAAGGCAAAATTAATCACTGAATTTCTTAAGTCACACAGGTGGAGT--- 2001
 DB 111 ValLeuLysAspGlyValThrLeuGluAlaLysLysValSerGlnThrAspGlySerThr 130
 QY 2002 CTGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACACACCAACACAGCCT 2061
 DB 131 ValValMetAspLeuGlyThr-----LeuGlnThr 141
 QY 2062 CTGCGCGTAAATCACTGATCAGGCTTCCCAATCTGCATTTGCTCTTCTTTCTTTGTTA 2121
 DB 142 SerSerSerGlyGluThrIleThrLeuThrAsnLeuAspIleAsnValAlaSer-----Leu 160
 QY 2122 GCAAAACAATGAGTTACGAATCTCTACCAATCTCCAGCGGCAAGATTTCTCATCTGCA 2181
 DB 161 GlyGlyGlyGlyValAlaProAspProAlaLysValGluAlaGlnAlaSer----- 177
 QY 2182 GTCATTGATGACCAACTGCTGGTCTCTTACAAATAGTGGCCCTATCTTTTGGAGAT 2241
 DB -----GlyLysThr-----ValThrIleAsnAla-----ValAsnLeuValAsp 190
 QY 2242 TTGGATGATCAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAAATCAATGTC 2301
 DB 191 ThrAspGlyAsnAlaThrGlu-----TyrProIleLeuAlaThrSerGlnProPheThrAla 209

QY 2302 CTGAAATTACAGTTAGGGACTAAGCCCCCAGCTAATGCCCATCAGATTGACTCTTAGGG 2361
 DB 210 IleIleAlaLysAlaGlySerGlyThrThrThrProThrAsp---AsnLeuLys 228
 QY 2362 AATGAGATGCT-----AAGTATGGCTATCAAGGAGCTGGAAGCTTGGTGGGATCCT 2415
 DB 229 AsnTyrThrProThrHisTyrGlyThrGlnGlyAsnThrThrValThrTrpLysLeu 248
 QY 2416 AATACAGCAAAATATGCTCTATATCTCTGAAAGCTACATCGGACTAAACCTGGGTATAAT 2475
 DB 249 GlyThrSerAlaGln---GluGluThrAlaThrLeuThrTrpGluGlnThrAspTyrSer 267
 QY 2476 CTGGGCTCTGAGGAGTAGCTCTTTGGTCCAAATAGTTATGGGATCACTATTTAGAT 2535
 DB 268 ProAsnProGluArgGlnGlyProLeuValProAsnThrLeuTrpGlySerPheSerAsp 287
 QY 2536 ATACGATCTCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTTGCGAGGA 2595
 DB 288 IleArgAlaIleGlnAsnLeuIleAspIleSerValAsnGlyAlaAspTyrArgArgGly 307
 QY 2596 TTATGGTCTTCTGAGTTCGAAATTTCTTCTATCATGACCGCATGCTTTAGGTACAGGA 2655
 DB 308 PheTrpValSerGlyLeuGlyAsnPheLeuHisLysSerGlySerAsnThrLysArgLys 327
 QY 2656 TATCGTATATTTAGTGGGTTTCTCTTAGGAGCAAACTCTTACTTTGGATCA---TCG 2712
 DB 328 PheArgHisSerAlaGlyTyrAlaLeuGlyValTyrAlaGlnThrSerThrGluAsp 347
 QY 2713 ATGTTGGTCTAGCATTCGGAATTTGGTATGATCTAAAGATTATGATAGTGTGCT 2772
 DB 348 ValPheSerAlaAlaPheCysGlnLeuPheGlyLysAspLysAspTyrPheValSerLys 367
 QY 2773 TCCAATCATGCTTCATAGGATCGCTTATCTA-----TCTACC 2814
 DB 368 AsnSerSerAsnIleTyrAlaGlySerIleTyrGlnHisIleSerTyrTrpAsnAla 387
 QY 2815 CAACAGCTTTATGTGATCTTATTTCTCGGAGATCGCTTATC-----CGT 2862
 DB 388 TrpGlnAsnLeuGlnSerThrIleGlyAlaGluAlaProLeuValLeuAsnAlaGln 407
 QY 2863 GCTAGCTACGGTTGGGAATCAGCATATCAAACTCATATACA----- 2907
 DB 408 LeuThrTyrCysHisAlaSerAsnAsnMetLysThrAsnMetThrAsnThrTyrValPro 427
 QY 2908 -----TTTGACAGGAGAGCGATGTTCTGGTGGATTAATACTCTCTGCTGGAGAG 2958
 DB 428 LysAsnValThrLeuSerGluIleLysGlyAspTrpGlyAsnAspCysPheGlyValGlu 447
 QY 2959 ATTGAGCGGATTTACCGATTGTGATTAATCTCAATCTAAGCTCTATTTGAATGATGCTGCT 3018
 DB 448 PheGlyAlaMetAlaProIleGlu---ThrProSerSerPheLeuPheAspArgTyrSer 466
 QY 3019 CTTTCTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAG---GAAGGC 3075
 DB 467 ProPheLeuGlnLeuGlnLeuValHisAlaHisGlnAspPheLysGluAsnAsnSer 486
 QY 3076 GATCAAGCTCGGCATTCAGAGCGGACATCTCTTAATATCTATCAGTTCTCTTGTGGAGTG 3135
 DB 487 AspGlnGlyArgTyrPheGluSerSerAsnLeuThrAsnLeuSerMetProIleGlyLe 506
 QY 3136 AAGTTTGAATGATGTTCTAGTACATCTCTTAATAATATAGCTTTATGGCGGCTTATATC 3195
 DB 507 LysPheGluArgPheAlaTyrAsnAspValAlaSerTyrHisLeuThrAlaAlaTyrAla 526
 QY 3196 TGTGATCTTATGCGACCATCTCTGGTACTGAGACAACGCTCTTA-----TCCCATCAA 3249
 DB 527 ProAspIleValArgSerAsnProAspCysThrAlaSerLeuLeuValSerProThrSer 546
 QY 3250 GAGACATGGCAACAGATGCTTTCATTATGACAGACATGGAGTTGTGGTTAGAGGATCT 3309
 DB 547 AlaValTrpValThrLysAlaAsnAsnLeuAlaArgSerAlaPheMetLeuGlnAlaGly 566


```
QY 1804 ATGGCAACGGAAATACACGACGCGAGCTCTCCAACTCTTAAATTAACGATGT 1863
Db 291 ValAlaGluGlyThr-----AlaAspSerAen-----LeuGluileAenLysAla 305
QY 1864 GAAGGA-----TACACAGGGGATATGTTTTGCT----- 1893
Db 306 AspGlyGlyThrSerTyThrGlySerilePheSerGlyArgTyriLeGluSerPro 325
QY 1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTAGCATAGCAAGGAAGG 1941
Db 326 HisLysArgMetLysHisValSerThrPheThrGlnProLeuThrLeuSerSerGlySer 345
QY 1942 ATTGTTCTTCGTGAAGCGCAAAATATCATGTAATCTCTAAGTCACAGAGTGGGAGT 2001
Db 346 LeuValLeuGluLysGlyAlaHisLeuLysAlaLysSerLeuThrGlnThrAlaGlySer 365
QY 2002 ---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCACACAG 2058
Db 366 LysValileLeuAaspGlnThrSerSerileGlu----- 376
QY 2059 CCTCTCGCGCTAAATCAGTTGATACGCTTTCCAACTGCAATTTGCTCTCTTTCTTTG 2118
Db 377 -----ThrLysGluAenileAaspileLysGluLeuTrrPLeuArgLeuAaspPhe 393
QY 2119 TTAGCAAAATGCAAGTTACGATCTCTCTACCAATCTCCAGCGCAAGATTTCTCATCT 2178
Db 394 -----AenThrProThr-----AlaThr 399
QY 2179 GCAGTCATTGTTAGCACAACTGCTGTTCTGTACAAATAGTGGGCTTATC---TTTTTT 2235
Db 400 CysIleSerThrSerGlyAenAlaHisThrileThrileLysGlyProLeuGlyValPhe 419
QY 2236 GAGGATTTGGATGATACAGCTTATGATAGGTAT-----GAT 2271
Db 420 ThrAasp---GlnGluThrPheTyrrAaspAenHisAlaLeuAlaTyrrSerileAaspGlnGlu 438
QY 2272 TGGCTAGGTCTTAATCAAAAATCAATCTCTGAAATTTACAGTTAGGACGTAAGCCCCCA 2331
Db 439 PheLeuGlnLeuAlaAaspLys---AaspileThrLysileSerLeuValAaspileProGln 457
QY 2332 GCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCCTAAGTAGTGGCTATCAAGGA 2391
Db 458 Ala-----ValArgLysAenLeuAaspSerHisSerGlyTyrrGlnGly 471
QY 2392 AGCTGGAAGCTTGGCTGGAT-----CCTAATACAGCAAAATATGTCCTTATACTCTG 2445
Db 472 LysTrpSerileAaspTrrLysThrValProGlySerThrAenAlaGlyValThrThrLeu 491
QY 2446 -----AAGCTACA-----TGGACTAAACTGGGTATATCTCT-----GGGCT 2484
Db 492 GlyThrLysThrAlaThrValHisTrpArgProThrGlyTrileProPheGlyGlySer 511
QY 2485 GAGCGAGTAGCTTCT---TTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATACGA 2541
Db 512 GlnGluileThrProLeuValValAenThrLeuTrpGlyAenPheSerAaspileArg 531
QY 2542 TCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGT---CGAGGATTA 2598
Db 532 AenLeuGluArgThrValGlu---SerLeuAlaValAenSerLeuCysSerGluGlyPhe 550
QY 2599 TGGGTTCTCGAGTTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGGTGAGGATAT 2658
Db 551 TrpAlaAlaGlyIleLysAenTyrrLysSerAenSerProAlaGluAenTyrrValPhe 570
QY 2659 CGGTATATTAGTGGGGTTATCTCTTAGGACGCAACTCTAC---TTTGGATCATCGATG 2715
Db 571 GlnHisHisAenAlaGlyTyrrAlaileGlyMetAenLysHisThrLeuSerGluAenVal 590
QY 2716 TTTGCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTTATGTAGTGTGCTGCTTCC 2775
Db 591 PheSerAlaAlaPheSerGlnLeuPheGlyLysAaspArgAaspHisAlaAenGlyHisVal 610
```

```
QY 2776 AATCATCATCTGCTGCATAGCATCGTTTAT-----CTATCTACCCAA 2817
Db 611 AspHisGlnThrLeuSerGlySerPheTyrrAlaHisHisValGlySerLeuPrometLeu 630
QY 2818 CAAGCTTTATGTGGATCTCTATTTGTTCCGAGATCGCTTTATCCGTGCT----- 2865
Db 631 ArgPheLeuCysGlyGlySerLysAenCysProProGluLeuGlnAlaSerProSerIle 650
QY 2866 -----AGCTACGGGTTTGGGAATCAGCATGATGAAACCTCA 2901
Db 651 ProValileValAenAlaGlnLeuSerTyrrSerHisSerAenAenHisLeuThrIleHis 670
QY 2902 TATACATTTCCAGAGGAGCGATGTTCTGTTGGATATAACTCTCTGCTGGTGGAGAGATT 2961
Db 671 HisGluAaspThrThrLysThrThrGlyMetTrpSerAenTyrrSerLeuAlaAlaGluLeu 690
QY 2962 GGACCGGGATTTACCGATTGTGATTACTTCCATCTAAGCTCTATTGTAATGAGTTCGGTCT 3021
Db 691 GlySerThrPheValTyrrThrLeuSerLysCysProSerileLeuLysAenValSerPro 710
QY 3022 TTCTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGATCAA 3081
Db 711 PheValLysLeuGlnGlyValTyrrSerGluGlnArgLysPheSerGluGlyLeuArg 730
QY 3082 GCTCGGCGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTT 3141
Db 731 ArgCysLeuPheSerSerThrTyrrLeuAlaAenLeuAlaLeuProLeuGlyIleLysIle 750
QY 3142 GATCGATGTTCTAGTACACATCTCTAAATATATAGCTTTATGGCGCTTATATCTGTGAT 3201
Db 751 HisGlyValCysProArgGluLeuPheAlaTyrrAaspLeuSerAlaMetTyrrValHisAasp 770
QY 3202 GCTTATCGCACCATCTCTGTTACTGACAGCAACG---CTCCTATCCCATCAAGAGACATGG 3258
Db 771 ValPheArgileAaspProGluThrMetThrLeuPheLeuileGlyGlyLeuAlaProTrp 790
QY 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTTGTTAGAGGATCTATGATGCT 3318
Db 791 ThrThrHisAlaAenAenLeuAlaThrLysAlaileValValGlnGlySerGlyArgPhe 810
QY 3319 TCTCTAACAGTATATAGAAATATATGGCCATATGAGATATGAGTATCGAGATGCTTCT 3378
Db 811 AlaValargSerAenileGluValPheAlaGluGlyAenCysGluLeuArgSerSerSer 830
QY 3379 CGAGCTTATGTTTGGAGTCGAGGAAGTAGAGTCGGTTC 3417
Db 831 HisSerTyrrAenTyrrAaspPheGlyAlaLysileHisPhe 843
```

Search completed: May 13, 2006, 10:31:16
Job time : 1252.5 sec

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 13, 2006, 09:33:09 ; Search time 123.2 Seconds
(without alignments)
4745.077 Million cell updates/sec

Title: US-10-701-844-1
Perfect score: 7883
Sequence: 1 gggcaaaactctcccccgcg.....gcactcaagaagaattc 4435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/abs/ABSWEB/spool/US10701844/runat_12052006_165417_26223/app_query.fasta_1
-DB=A.Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=blos62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p
-USER=US10701844 @CGN 1.1 348 @runat_12052006_165417_26223 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	66.8	1012	2	AAY16735 C. tracho
2	5131.5	65.1	1013	5	ABG91021 Chlamydia
3	5131.5	65.1	1013	7	ADD43798 Chlamydia
4	5131.5	65.1	1013	7	ADD43722 Chlamydia
5	5131.5	65.1	1013	7	ADD43718 Chlamydia
6	5131.5	65.1	1013	7	ADD43708 Chlamydia
7	5131.5	65.1	1013	7	ADD43714 Chlamydia
8	5131.5	65.1	1013	7	ADD43786 Chlamydia
9	5131.5	65.1	1013	7	ADD43788 Chlamydia

10	5131.5	65.1	1013	7	ADD43710	Chlamydia
11	5131.5	65.1	1013	7	ADD43706	Chlamydia
12	5131.5	65.1	1013	7	ADW29027	C. trachom
13	5131.5	65.1	1013	9	AEA19078	Chlamydia
14	5131.5	65.1	1013	9	AEA19002	Chlamydia
15	5125.5	65.0	1013	9	AEA19080	Chlamydia
16	5123.5	65.0	1013	9	AEA19006	Chlamydia
17	5123.5	65.0	1013	9	AEA19090	Chlamydia
18	5123.5	65.0	1013	9	AEA19010	Chlamydia
19	5121.5	65.0	1013	9	AEA19000	Chlamydia
20	5121.5	65.0	1013	9	AEA18998	Chlamydia
21	5112.5	64.9	1013	9	AEA19014	Chlamydia
22	5106.5	64.8	1013	2	AA16737	C. tracho
23	5106.5	64.8	1013	2	AA16738	C. tracho
24	5090	64.6	1006	4	AAG83207	Protein e
25	5090	64.6	1006	5	AB94178	Chlamydia
26	5084	64.5	982	3	AB13633	C. tracho
27	5084	64.5	982	3	AAG83201	Protein e
28	5084	64.5	982	5	AB94172	Chlamydia
29	5083	64.5	1006	3	AB13639	C. tracho
30	3336.5	42.3	670	7	ADD42756	Chlamydia
31	2606	33.1	524	2	AA16738	Chlamydia
32	2547	32.3	505	2	AA16739	C. tracho
33	2350	29.8	458	2	AA16751	Chlamydia
34	1735	22.0	325	2	AA16752	Chlamydia
35	1715	21.8	708	2	AA16753	Chlamydia
36	1590.5	20.2	631	4	AAG83274	Chlamydia
37	1590.5	20.2	631	5	AB94245	Chlamydia
38	1589	20.2	1016	4	AA16751	Chlamydia
39	1589	20.2	1016	5	ABG91037	Chlamydia
40	1589	20.2	1016	7	ADD42682	Chlamydia
41	1589	20.2	1016	7	ADD43800	Chlamydia
42	1589	20.2	1016	9	ADW29028	C. trachom
43	1589	20.2	1016	9	AEA19092	Chlamydia
44	1456	18.5	664	7	ADD42755	Chlamydia
45	1365.5	17.3	973	5	AB90527	Chlamydia

ALIGNMENTS

RESULT 1
AAY16735
ID AAY16735 standard; protein; 1012 AA.
XX
AC AAY16735;
XX
DT 21-JUL-1999 (first entry)
XX
DE C. trachomatis LGV L2 HMW protein.
XX

Chlamydia; high molecular weight protein; HMW protein; urethritis;
bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
KW salpingitis; tubal occlusion; infertility; cervical cancer;
arteriosclerosis; atherosclerosis.
XX
OS Chlamydia trachomatis.
XX
PN WO917741-A1.
XX
PD 15-APR-1999.
XX
PP 01-OCT-1998; 98WO-US020737.
XX
PR 02-OCT-1997; 97US-00942596.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson JW, Pace JL;
XX WPI; 1999-287659/24.
XX N-PSDB; AAX60539.

PT New Chlamydia protein useful for treating conjunctivitis, urethritis and
 PT cervical cancer.

XX Claim 4; Fig 3; 141pp; English.

XX The invention relates to an isolated Chlamydia species high molecular
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
 CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
 CC be used for preventing, treating or ameliorating a disorder related to
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products
 CC can also be used for detection and diagnosis. The present sequence
 CC represents a C. trachomatis HMW protein

XX Sequence 1012 AA;

Alignment Scores:

Pred. No.:	0	Length:	1012
Score:	5267.00	Matches:	1012
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	66.8%	Indels:	0
DB:	2	Gaps:	0

US-10-701-844-1 (1-4435) x RAY16735 (1-1012)

QY	302	ATGCAAGCTCTTTCATAGTCTCTTCAATGATTCTAGCTATTCTTGCTGCTCT	441
DB	1	MetGlnThrSerPheHisLeuSerMetIleLeuAlaIleSerCysSer	20
QY	442	TTAAATGGGGGATATGCGACGAATCATGTTCTCCTCAAGGAATTTACGATGGGAG	501
DB	21	LeuAsnGlyGlyTyAlaAlaGluIleMetValProGlnGlyIleTyAspGlyGlu	40
QY	502	ACGTTAAGTATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTCTGTTTTT	561
DB	41	ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGACGAGAGTTAAACATTAATAATCTTGCAATCTTATGACGCTTTCGCTTTAAGT	621
DB	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer	80
QY	622	TGTTTGGGAATTTATAGGAGTTTACTGTGTTTGGGAGAGGACACTCGTTGACTTTC	681
DB	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
QY	682	GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGGCTGCTGATGGACTG	741
DB	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu	120
QY	742	TTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAATTGCAATTCATTCTCCGCTA	801
DB	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACACACATCTACACCGTCT	861
DB	141	LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGTAATATTATTTCTAAACAGACTTTTGTGTTTACTCAATTAATGAGAGTCTCATTC	921
DB	161	AsnGlyThrIleTySerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe	180
QY	922	TATAGTAATATTAGTCTGGAGATGGGAGCTATAGATGCTTAAGCTTAACGGTTCAA	981
DB	181	TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAGCTTTGTGCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTGT	1041
DB	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys	220
QY	1042	CAAGTAGTACCAGTTCTCTGCTATGCTTAACGAGGCTCTATTGCGCTTTGTAGCGAAT	1101

DB	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn	240
QY	1102	GTTCCAGGAGTAAGAGGGGAGGATTGCTGCTCTTTCAGGATGGGCAGCAGGAGTGTCA	1161
DB	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCCGCTAGAGTTT	1221
DB	261	SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATCGGACGTAGCCCGCAGTAGGAGGAGGATTCTCTACGGGAAAGTGTCTTCTCTG	1281
DB	281	AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu	300
QY	1282	AATAATGAAAAACCTTTGTTCTCACAATGTTCTCTCTCTCTCTCTCTCTCTCTCTAAG	1341
DB	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyIleAlaAlaLys	320
QY	1342	CAACCAACAGTGCACAGGCTTCTTAATACGAGTAAATTAATACGAGATGGAGGAGCTATC	1401
DB	321	GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyGlyAspGlyAlaIle	340
QY	1402	TTCTGTAAGATGTCGCGCAGCAGGATCCCAATTAATCTCTGATCAGTTTCTCTTGCATG	1461
DB	341	PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly	360
QY	1462	GAGGAGTAGTTTCTTTTAGTAGCAATGATGCTCTGCGAAAGGGGAGCTATTATATGCC	1521
DB	361	GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyAla	380
QY	1522	AAAAAGCTCTCGTGTGCTAATCTGCGCCCTCAATTTTAAAGAAATATCGCTAATGAT	1581
DB	381	LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp	400
QY	1582	GCTGCGAGATTTTATAGGAGATCTGGAGACTCTGAGCTCTGTTTCTCTGATTTATGGAG	1641
DB	401	GlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyroGlyAsp	420
QY	1642	ATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAGAATGCTGCCGATGTTAATGCG	1701
DB	421	IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly	440
QY	1702	GTAATGTCCTCACAAGCCATTTGATGGATCGGAGGCGGAGGAAATTAACGACATTAGA	1761
DB	441	ValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIleThrThrLeuArg	460
QY	1762	GCTAAGCAGGCGCATCAGATCTCTTTAATCATCCCATCGAGATCGCAACCGAAATAAC	1821
DB	461	AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn	480
QY	1822	CAGCAGCGCAGTCTTCCAAATCTTAAAAATTAAACGATGGTGAAGGATACACAGGGAT	1881
DB	481	GlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGlyTyThrGlyAsp	500
QY	1882	ATTGTTTTGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGAGG	1941
DB	501	IleValPheAlaAsnGlySerSerThrLeuTyGlnAsnValThrIleGluGlnGlyArg	520
QY	1942	ATTGTTCTCTGAAAGGCAAAATATCAGTGAATCTCTAAGTCACAGCAGGTCGGAGT	2001
DB	521	IleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySer	540
QY	2002	CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACACAGCTC	2061
DB	541	LeuTyMetGluAlaGlySerThrTrpAspPheValThrProGlnProGlnGlnPro	560
QY	2062	CCTGCCCTAATCAGTTGATCAAGCTTTCCAAATCTGCAATTTGCTCTCTTTCTTTGTTA	2121
DB	561	ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeu	580
QY	2122	GCMAACATGAGTTACGAATCCCTACCAATCTCTCCAGCGCAAGATTTCTCATCTGCA	2181

581 AlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisProAla 600
2182 GTCATTGGTAGCACAACCTGCTGCTGTTTACAAATAGTAGGGCCCTATCTTTTTCAGGAT 2241
601 ValIleGlySerThrAlaGlySerValThrIleSerGlyProIlePheGluAsp 620
2242 TTGGATGATACAGCTTATGATAGGTATGATCGCTAGGTTCCTAAATCAAAATCAATGTC 2301
621 LeuAspAspThrAlaTyrrAspArgTyrrAspTrpLeuGlySerAsnGlnIleAsnVal 640
2302 CTGAATATACAGTTAGGGACTAAGCCCCAGCTAAATGCCCCATCAGATTGTAGCTTAGGG 2361
641 LeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGly 660
2362 AATGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCGGATCCTAATACA 2421
661 AsnGluMetProLysTyrrGlyTyrrGlnGlySerTrpLysLeuAlaTrpAspProAsnThr 680
2422 GCAAAATAATGGTCTTATCTCTGAAAGCTACATGAGCTAAATCTGGGTATATATCTGGG 2481
681 AlaAsnAsnGlyProTyrrThrLeuLysAlaThrTrpThrLysThrGlyTyrrAsnProGly 700
2482 CCTGAGCGAGTAGCTCTTTGGTTCCAAATAGTTATGGGATCCATTTTATAGATACGA 2541
701 ProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArg 720
2542 TCTGGCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTTGTCGAGATTATGG 2601
721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrrCysArgGlyLeuTrp 740
2602 GTTCTGGAGTTTCGAATTTCTTATCATGACCGGATGCTTTAGTTCAGGAGATATCGG 2661
741 ValSerGlyValSerAsnPhePheTyrrHisAspArgAspAlaLeuGlyGlnGlyTyrrArg 760
2662 TATATTAGTGGGGTTATTTCTTAGGAGCAAACTCTACTTTGGATCATCGATGTTTGGT 2721
761 TyrrIleSerGlyGlyTyrrSerLeuGlyAlaAsnSerTyrrPheGlySerSerMetPheGly 780
2722 CTAGCATTTACGAAGTATTTGTAGATCTAAAGATTATGTAGTGTGCTTCCATCAT 2781
781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrrValValCysArgSerAsnHis 800
2782 CATGTTGTGATAGTACCGTTTATCTTATCTACCAACAGCTTATGTGATCTCTATTGG 2841
801 HisAlaCysIleGlySerValTyrrLeuSerThrGlnGlnAlaLeuGlySerTyrrLeu 820
2842 TTCGAGATGCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCA 2901
821 PheGlyAspAlaPheIleArgAlaSerTyrrGlyPheGlyAsnGlnHisMetLysThrSer 840
2902 TATACATTTGCAGAGAGCGATGTTCTGTTGGGATTAATCACTGCTGGCTGGAGAGATT 2961
841 TyrrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysAlaLeuAlaGlyGlu 860
2962 GGAAGCGGATTCAGGATGATGATCTCCATCTCAAGCTTATTTGAATGATGTTGCGTCT 3021
861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrrLeuAsnGluLeuArgPro 880
3022 TTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGGATCAA 3081
881 PheValGlnAlaGluPheSerTyrrAlaAspHisGluSerPheThrGluGluGlyAspGln 900
3082 GCTCGGCGATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGTGGAGTGAATTT 3141
901 AlaArgAlaPheLysSerGlyHisLeuLeuAsnLeuSerValProValGlyValLysPhe 920
3142 GATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGGCGCTTATATCTGTGAT 3201
921 AspArgCysSerSerThrHisProAsnLysTyrrSerPheMetAlaAlaTyrrIleCysAsp 940
3202 GCTTATCGCACCATCTCTGGTACTGACAGCAACGCTCTCTATCCCATCAAGAGACATGGACA 3261
941 AlaTyrrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThr 960

3262 ACAGATGCTTTTCATTATTAGCAAGCATGGAGTTGGTTAGAGGATCTATGTATGCTTCT 3321
961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrrAlaSer 980
3322 CTAACAAGTAATATAGAAAGTATATGCGCATGCGCAATGGAAGATATGAGATCGAGATCCTTCTCGA 3381
981 LeuThrSerAsnIleGluValTyrrGlyHisGlyArgTyrrGluTyrrArgAspAlaSerArg 1000
3382 GCGTATGCTTTTCAGTGCAGGAGTAGATGTCGGTTTC 3417
1001 GlyTyrrGlyLeuSerAlaGlySerArgValArgPhe 1012

RESULT 2

ABG91021
ID ABG91021 standard; protein; 1013 AA.

AC ABG91021;

DT 29-NOV-2002 (first entry)

DE Chlamydia trachomatis outer membrane protein G protein.

KW Gram-negative bacterial bleb; PorB; outer membrane protein;
Chlamydia trachomatis infection; Chlamydia pneumoniae infection;
protective antigen; antibacterial; vaccine; gene; ds.

OS Chlamydia trachomatis.

PN WO200262380-A2.

PD 15-AUG-2002.

PF 08-FEB-2002; 2002WO-EP001356.

PR 08-FEB-2001; 2001GB-00003169.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;

DR WPI; 2002-657510/70.

N-PSDB; ABS67342.

PT Novel gram-negative bacterial bleb presenting on its surface PorB outer
membrane protein from Chlamydia trachomatis or protective antigen from
Chlamydia pneumoniae, useful for preventing Chlamydia infection.

PS Disclosure; Page 10; 75pp; English.

CC The present invention relates to a new gram-negative bacterial bleb
presenting on its surface the PorB outer membrane protein from Chlamydia
trachomatis, or a protective antigen from C. pneumoniae. The invention is
useful for preventing C. trachomatis or C. pneumoniae infection in a
host. The present nucleic acid sequence represents a Chlamydia
trachomatis gene of the invention

SQ Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: Gaps: 1

US-10-701-844-1 (1-4435) x ABG91021 (1-1013)

382 ATCAACAGCTTTTCCATAGTCTTCTTTCATGATTCAGTATTATCTTCTGCTCTCT 441

1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrrSerCysSer 20

QY 442 TTAATGGGGGGATATGACGACGAATCATGTTCTCTCAAGCAATTTACGATGGGAG 501
Db : : : : :
QY 21 LeuSerGlyGlyTyAlaAlaGluIleMetIleProGlnGlyIleTyAspGlyGlu 40
QY 502 ACGTTAACTGATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTT 561
Db 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrValPhe 60
QY 562 TCTCGAGGAGGTAAACATTAATAAATCTTGACAAATCTATGACGCTTCCCTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTGGGAACCTATTAGGAGGTTTACTGTTTGGGAGGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACCGACTTCTACAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTACTATTGAGGGTTTTAAAGAATTATCTTTTCCAAATTCGAATTCATTACTTCCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCACGACTAATAAGGGTAGCGAGCTCCGACGACGACATCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGTACTATTATTCTAAACACAGACTTTTGTGTACTCAATAATGAGAAGTTCTCATTC 921
Db 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTGCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys 220
QY 1042 CAAGTAGTCACCAAGTTCTTGCTATGCTTAACGAGGCTCTTATGCTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
QY 1102 GTTCAGGAGTAAGAGGGGAGGATGCTGCTGCTCAGGATGGCAGCGAGGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAAGTAGCCGAGTAGGAGGATTTACTCTACGGAACGTTGCTTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAACCTGTGTTCTCAACATGTGCTTCTCTGTTTACATTCCTCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyIleAlaAlaGlu 320
QY 1342 CAACCAAGTAGGAGCGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTGAAGATGGTGGCAA--GCAGGATCCAAATCACTCGGATCAGTTTCTTTGAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGGAGGAGTAGTTTCTTTTATAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTy 380
QY 1519 GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578

Db 381 AlalysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATCGTGGAGCGATTTATTTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638
Db 401 AspGlyGlyAlaIleTyLeuGlyLeuSerGlyGluLeuSerLeuSerAlaAspTyGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTTAAAGAACACCCAAAGAAATGCTGCCGATGTTAAT 1698
Db 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GGCGTAATCTGCTCTCACAGCCATTTCCGATCGGATCGGGAGGAGGAGGAGGAGGAGGAGG 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACGAGCAGGCGAGTCTTCCAACTTCTTAAATAATTAAACGATGCTGAAGGATACACAGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly 500
QY 1879 GATATTGTTTGTCTTAATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGACAGAGA 1938
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTGTTCTTCGTAAGGCGAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACAG 2058
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTGCGCTTAATCAGTTCATCAGCTTTCCAACTGCAATTTGTCTTCTTCTTCTTCTG 2118
Db 561 ProProAlaAlaAsnGlnIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAACAATCAGATTACGAATCTCTCTACCAATCTCCAGCGCAAGATTCATCCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 CGAGTCATGTTAGCACAACCTGCTGCTGTTTCTTACAAATAGTGGCCCTATCTTTTTCAG 2238
Db 601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTGAAAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGATGAGATGCCCTAAGTATGGCTATCAAGGAGCTGGAAGCTTGGGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAAAATATGTCCTTATCTCTGAAAGCTACATCGACTTAAACCTGGGTATATCCT 2478
Db 681 ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTrpLysThrGlyTyAsnPro 700
QY 2479 GGGCTGAGCAGTAGTACTTCTTTGTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCATTTCAGCAATTCAGCAAGTCTGGATGGCGCTCTTATGTCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGATTTCGAATTTCTTCTATCATGACCGCGATGCTTTTAGTCAGGATAT 2658

```
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGAGCAAACTCCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTATACCCAAAGCTTTATGTGCATCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTCCGAGATCGGTTTATCCGCTAGCTACGGTTTGGGAATCAGCATATGAACC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2899 TCATATACATTTGCAGAGAGCGATGTTCTGTGGATAATAACTGTCTGCTGGAGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGGAGCGGGATTACCGATTGTGATTACTCCATCTAAGCTCTTATTTGAATGAGTTCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy 3019 CTTTCTGTCGAAGCTGAGTTTCTTATGCGCATCATGAATCTTTACAGAGAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGlyAsp 900
Qy 3079 CAGCTCGGGCTTCAAGAGCGACATCTCTAAATCTATCATGTTCTCTGTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCGACCACTCTCTGTACTGTAGACACGCTCTCTATCCATCAGAGATCG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATCGCTTTCATTTAGCAAGACATGGAATGTTGTGTAGAGATCTATGTATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy 3319 TCTCTAACCAAGTATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGGCTATGTTTGTAGTGAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 3
ADD43798
ID ADD43798 standard; protein; 1013 AA.
AC ADD43798;
XX
XX
XX 15-JAN-2004 (first entry)
DT
DE Chlamydia trachomatis immunogenic protein, SEQ ID No 93.
XX
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO2003049762-A2.
XX
PD 19-JUN-2003.
XX
```

```
PF 12-DEC-2002; 2002WO-IB005761.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Ratti G;
XX
DR WPI; 2003-532882/50.
DR N-PSDB; ADD43799.
XX
PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
XX
PS Claim 6; SEQ ID NO 93; 164pp; English.
XX
XX The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of
CC the 131 C. trachomatis proteins with immunogenic properties of the
CC invention.
XX
SQ Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: Gaps: 1

US-10-701-844-1 (1-4435) x ADD43798 (1-1013)
Qy 382 ATGCAAACTCTTTCCATAAGTTCTTTCTTTCAATGATTTAGCTTATTCTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCys 20
Qy 442 TTAATATGGGGGGATATGACAGAAATCATGGTTCCTCAAGAAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
Qy 502 AGTTAACTGTATCATTTCCCTATCTCTATAGGAGATCCGAGTGGACTGTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTGCAGGAGGTTAAACATTTAAATAATCTTGACAAATCTTATTCAGCTTTTAAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
Qy 622 TGTTTTGGGAACCTTATTAGGAGCTTTTACTGTGTTTAGGGAGAGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACGAGCTTCTACAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
```

QY 742 TTTACTATTGAGGGTTTAAAGATTATCTCTTTTCCAAATTCGAATTCATTACTTCCCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGAGCTAATAAGGGTAGCCAGACTCCGACGACAACTATCACCGCTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTTAAACAGATCTTTTGTCTACTCAATAATGAGAAGTCTTCATTC 921
Db 161 AenGlyThrIleTySerLysThrAspLeuLeuLeuLeuAenAenGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA 981
Db 181 TyrSerAenLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTACCAAGCTTTGTCTCCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuLysValPheGlnGlnAenThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTCACCAAGTTTCTCTGCTATAGCTTAAACGAGGCTCTTATTCCTTTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240
QY 1102 GTTGCAAGAGTAAGAGGGGGAGGATGCTGCTGTTCAAGATGGCCAGCAGGAGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaLysValGlnAenGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTAGCCGAGTAGGAGGAGTTTACTCTCTACGGGAACGTTCCTTCTCG 1281
Db 281 AspGlyAenValAlaArgValGlyGlyIleTySerTyGlyAenValAlaPheLeu 300
QY 1282 AATAATGGAACCTTGTCTTCTCAACATGTTGCTTCTCTGTTTACATTCGCTCTAAG 1341
Db 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyIleAlaAlaGlu 320
QY 1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAenThrSerAspAenTyGlyAspGlyAlaIle 340
QY 1402 TTCTGTAAAGATGTGCGCAA---GCAGGATCCAAATAACTCTGGATCAGTTTCTTGTAT 1458
Db 341 PheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAsp 360
QY 1459 GGAGAGGGATGATGTTTCTTTAGTAGCAATGATGCTGCGGAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyAlaIleTy 380
QY 1519 GCCAAAAGCTCTCGGTGCTAAGTGTGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
QY 1579 GATGGTGGAGCATTTATTTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTATGA 1638
Db 401 AspGlyGlyAlaIleTyLeuGlyGlyLysGlyGlyLeuSerLeuSerAlaAspTyGly 420
QY 1639 GATATTTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTCCGATGTTAAT 1698
Db 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440
QY 1699 GCGGTAACTGTCTCTCAAGCCATTTCTGATGGGATCGGAGGAGGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrLeu 460
QY 1759 AGAGCTAAAGCAGGAGCATCTCTTTTAAATGATCCCATCGAGATGCCAAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAspProIleGluMetAlaAenGlyAen 480
QY 1819 ACCAGCAGGCGAGCTCTCCAAACTTCTTAAAAATTAACGATGGTGAAGGATACACAGG 1878

Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGluGlyTyThrGly 500
QY 1879 GATATGTTTGTCTTAATGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyGlnAenValThrIleGluGlnGly 520
QY 1939 AGGATGTTCTTCTGTAAGGCAAAATTTATCAGTGAATTCCTTAAGTCACACAGGTGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACCAACCAACAG 2058
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCCTCGCGTAATAGTTGATCACGTTTCCCAATCTGCAATTTGCTCTTCTCTTTG 2118
Db 561 ProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAACAATGCAAGTTACGAATCCTCTTACCAATCCTCCAGCGCAAGATTCATCCT 2178
Db 581 LeuAlaAenAenAlaValThrAenProProThrAenProAlaGlnAenSerHisPro 600
QY 2179 GCAGTCAATTGCTAGCACAACTGCTGTTCTTACAATTAGTGGGCTTATCTTTTTCAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATTTGGTAGTCTCTTAATCAAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAenGlnLysIleAsp 640
QY 2299 GTCTGAAATTTACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660
QY 2359 GGGATGAGATGCTTAACTATGGCTATCAAGAAAGCTGGAAGCTTGGCGGATCCTAAT 2418
Db 661 GlyAenGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY 2419 ACAGCAATAATGTGCTTATATCTCTGAAGCTACATGACTAAACCTGGGTATTAATCCT 2478
Db 681 ThrAlaAenAenGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAenPro 700
QY 2479 GGGCTGAGCAGATGCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTACATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGTCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGTTTCAAAATTTCTTATCATGACCGCGATGCTTTTAGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAenPheTyHisAspArgAspAlaLeuGlyGlnGlyTy 760
QY 2659 CGGTATATTAGTGGGGTTATCTTACGAGCAAACTCTTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAenSerTyPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTTACCAGAGTATTTGCTAGATCTTAAAGATTAATGATGTGTGTTCCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAen 800
QY 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCCAAACAGCTTTTATGTGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820
QY 2839 TTGTTCCGAGATCGCTTTATCCGCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyGlyPheGlyAenGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGGAGCATGTTCTGTCGGGATAATAACTGTCTGGCTGCGAGAG 2958

Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
 Qy 2959 ATTGGAGCGGATTACCATCTGATTAATCCATCAAGCTTATTTCAATGAGTTGGT 3018
 Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
 Qy 3019 CTTTTCGTCGACGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGAGCGCAT 3078
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
 Qy 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTTAATCTATCATGTTCTTGGATGAG 3138
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
 Qy 3139 TTTGATCCAGTCTTACATACATCTTAATATAGCTTTATGCGGCTTATATCTGT 3198
 Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
 Qy 3199 GATGCTTATCGCACCATCTCTGCTACTGAGACAGCGCTCTATCCCATCAAGAGACATGG 3258
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
 Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGTGTGGTTAGAGGATCTATGATGCT 3318
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 Qy 3319 TCTCTAACAGTAATATAGATATATGCGCATGATGCGGATGATGATGATGATGATGCTTCT 3378
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 Qy 3379 CGAGCTATGTTTTCAGTGCAGGAGTAGAGTCCCGTTC 3417
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 4

ADD43722
 ID ADD43722 standard; protein; 1013 AA.
 XX
 AC ADD43722;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 17.
 XX
 KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
 KW gene therapy; antibacterial.
 XX
 OS Chlamydia trachomatis.
 XX
 PN WO2003049762-A2.
 XX
 PD 19-JUN-2003.
 XX
 PF 12-DEC-2002; 2002WO-IB005761.
 XX
 PR 12-DEC-2001; 2001GB-00029732.
 PR 06-AUG-2002; 2002GB-00018233.
 PR 14-AUG-2002; 2002GB-00018924.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Grandi G, Ratti G;
 XX
 DR WPI; 2003-532882/50.
 DR N-PSDB; ADD43723.
 XX
 PT New immunogenic composition having a protein or encoding nucleic acid,
 PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
 PT infection.
 XX
 PS Claim 6; SEQ ID NO 17; 164pp; English.
 XX
 CC The invention relates to a novel immunogenic composition comprising a

CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
 CC acid comprises any of 131 fully defined amino acid or nucleotide
 CC sequences given in the specification, or has 50% or greater sequence
 CC identity to it, or their fragments. The protein and/or nucleic acid of
 CC the immunogenic composition is useful in the manufacture of a medicament
 CC for the treatment or prevention of infection due to Chlamydia
 CC trachomatis. The infection is treated or prevented by the medicament
 CC eliciting an immune response which is specific to a C. trachomatis
 CC elementary body, or for neutralising C. trachomatis elementary bodies,
 CC hence the immunogenic composition can be used in creating a vaccine. The
 CC immunogenic compositions can also be used for the diagnosis of C.
 CC trachomatis infection. The nucleic acids of the immunogenic compositions
 CC can be used to treat disorders by gene therapy. The immunogenic
 CC compositions have antibacterial activity. This sequence represents one of
 CC the 131 C. trachomatis proteins with immunogenic properties of the
 CC invention.

XX SQ Sequence 1013 AA;

Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5131.50 Matches: 983
 Percent Similarity: 98.8% Conservative: 18
 Best Local Similarity: 97.0% Mismatches: 11
 Query Match: 65.1% Indels: 1
 Gaps: 1
 DB:

US-10-701-844-1 (1-4435) x ADD43722 (1-1013)

Qy 382 ATGCAAAAGCTCTTTCCATTAAGTTCTTTCTTCAATGATCTAGCTTATTTCTGCTCTCT 441
 Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
 Qy 442 TTAATGGGGGATATGCGACAGAAATCATGGTTCTCAAGGATTTACCATGGGAG 501
 Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
 Qy 502 ACGTTAACTGTATCATTTCCCTATATCTGTATAGAGATCCGAGTGGGACTACTGTTT 561
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 Qy 562 TCTGAGGAGAGTTAAACATTAATAAATCTTGAATTTCTATTTGCGAGCTTTGCTTAAAGT 621
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
 Qy 622 TGTTCGGAAGTATTAGGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 Qy 682 GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTCTGATGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 Qy 742 TTTACTATTAGGGTTTAAAGAAATATCTCTTTTCAATGCAATTCATTTACTTCCGCTA 801
 Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
 Qy 802 CTGCTGCTGCAACCACTAATAAGGTAGCCAGACTCCGACGACACATCTACACCGTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
 Qy 862 AATGGTACTATTATTCTAAAAACAGATCTTTTGTACTCAATAAGAGAGTTCTCATTT 921
 Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
 Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
 Qy 982 GGAATTAGCAAGCTTTGTGTTCTTCCAGAAATATCTGCTCAAGCTGATGGGAGCTTCT 1041
 Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220

1042 CAAGTAGTCACCAAGTTCTCTGCTATGCTAAACGAGGCTCCTATTGCTTTTGTAGCGAAT 1101
Db
221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy
1102 GTTGCAGAGTAAGAGGGGAGGAGTTCGTCTCTTTCAGATGGCGAGGAGGTGTCA 1161
Db
241 ValAlaGlyValAlaGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
Qy
1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTCCGAAATATCTGCGGTAGAGTTT 1221
Db
261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Qy
1222 GATGGGACGTAGCCAGTAGGAGGAGGATTTACTCTCTACGGCAACGTTGCTTTCTCTG 1281
Db
281 AspGlyAsnValAlaAlaGValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Qy
1282 AATAATGGAAAACTTTGTTCTCAACAAATGTTCTCTCTCTGTTTACATTGCTGCTAAG 1341
Db
301 AsnAsnGlyIleThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
Qy
1342 CAACCAACAGTGCACAGGCTTCTAATACGAGTAATAATACGGAGATGGAGGAGCTATC 1401
Db
321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
Qy
1402 TTCTGTAGAATGGTGCACAA---GCAGGATCCCAATTAATCTCTGGATCAGTTTCTCTTGAT 1458
Db
341 PheCysIleAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
Qy
1459 GGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGGAAAGGGGAGCTATTAT 1518
Db
361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIleGlyAlaIleTyr 380
Qy
1519 GCCAAAAAGCTCTGGTTGCTAACTGTGGCTGTACAAATTTTAAGCAATATCGCTAAT 1578
Db
381 AlaIleLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
Qy
1579 GATGTCAGGAGCATTTATTAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
Db
401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420
Qy
1639 GATATTATTTTCGATGGGAATCTTAAAAAGACAGCAAGAGAGATGTCGCCGATTTAAT 1698
Db
421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
Qy
1699 GCGCTAATCTGCTCTCACAGCCATTTTCGATGGGATCGGAGGAGAAATACGACATTA 1758
Db
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyIleThrThrLeu 460
Qy
1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818
Db
461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
Qy
1819 AACACGCCAGCCAGCTTTCCAAACTTTAAAAATTAACGATGGTGAGGATACACAGG 1878
Db
481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyIleTyrThrGly 500
Qy
1879 GATATTGTTTTCCTAATGAGACGATCTTTGTACCAAAATGTTACGATAGCAAGCA 1938
Db
501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Qy
1939 AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998
Db
521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Qy
1999 AGTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACACACACAG 2058
Db
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
Qy
2059 CTCTCTGCGCTAATCAGTTGATCAGCTTCCCAATCTGATTTGCTCTTTCTCTTTG 2118
Db
561 ProProAlaAlaAsnGlnIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
Qy
2119 TTAGCAACAACATGCACTAGTACCAATCTCTTACCAATCTCTCAGCGCAAGATTCTCATCT 2178

581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
Qy
2179 GCAGTCATTCGTAGCACAACTGCTGTTCTGTTCACAAATTAGTGGCCCTATCTTTTTCAG 2238
Db
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy
2239 GATTTCGATGATACAGCTTATGATAGTATGATGGCTAGTCTTAATCAAAAAATCAAT 2298
Db
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
Qy
2299 GTCTCGAAATACAGTTAGGAGTAAGCCCCAGCTAATGCCCATCAGATTTTTCACCTA 2358
Db
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Qy
2359 GGGAAATCAGATCGCTAAGTATGGCTATCAAGAGAGCTGGAAGCTTGGCTGGGATCTTAAT 2418
Db
661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
Qy
2419 ACAGCAAAATATGCTCCTTATCTCTCAAAAGCTACATGAGCTAAAACTGGGTATAATCCT 2478
Db
681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
Qy
2479 GGGCTGAGCGAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATATA 2538
Db
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
Qy
2539 CGATCTCGCAATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTATGTCGAGGATTA 2598
Db
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy
2599 TGGGTTTCTGAGGATTCGAATTTCTTCATCATGACCGCGATGCTTTTAGTTCAGGGATAT 2658
Db
741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy
2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAACTCTACTTTGGATCATCGATGTTT 2718
Db
761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Qy
2719 GGTCTAGCATTTACCGAAGTATTTGTTAGATCTAAAGATATTAGTAGTGTGCTGCTCCAAAT 2778
Db
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy
2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCCCAACAGCTTTATGCGATCTCTAT 2838
Db
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy
2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898
Db
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy
2899 TCATATACATTTGACAGAGAGAGCGATGTTTCGTTCCGATTAATAACTGTCTGCTCGAGAG 2958
Db
841 SerTyrThrPheAlaGlnGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
Qy
2959 ATTGGAGCGGATTAACGATTCGATTCATCTAAGCTCTATTGATGATGATGCGCT 3018
Db
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy
3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGATCTTTTACAGGAGGCGCAT 3078
Db
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy
3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTCAAAATCTATCAGTTCTCTGTTGAGTGAAG 3138
Db
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy
3139 TTTGATTCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGCGCTTATCTGT 3198
Db
921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy
3199 GATGCTTATCGCACCATCTCTGCTAGACAGCGCTCTCTATCCATCAAGAGACATGG 3258

Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTAGACAGACATGAGAGTTGGTTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACCAAGTAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGGTTTGGTTCAGGAGAGTATAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 5

ADD43718
ID ADD43718 standard; protein; 1013 AA.
XX
AC ADD43718;
XX
DT 15-JAN-2004 (first entry)
XX
DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 13.
XX
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO2003049762-A2.
XX
XX 19-JUN-2003.
XX
PD 12-DEC-2002; 2002WO-IB005761.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2001; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Grandi G, Ratti G;
XX
XX WPI; 2003-532882/50.
DR N-PSDB; ADD43719.
XX
XX New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
XX
XX Claim 6; SEQ ID NO 13; 164pp; English.
PS
PS The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of
CC the 131 C. trachomatis proteins with immunogenic properties of the
XX invention.
XX
XX Sequence 1013 AA;
SQ

Alignment Scores:

Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1
US-10-701-844-1 (1-4435) x ADD43718 (1-1013)
QY 382 ATGCAAAAGCTCTTTCCATAAGTCTCTTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGGATATGACAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGTATCATTTCCCTATCTGTATAGGAGATCCGAGTGGGACTACTGTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGAGTTAACTTAAATAATCTTGCAATTTCTATTGACGCTTTGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTTGGAACTTATTAGGGAGTTTTACTGTTTTAGGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGAGCTTCTACAAATGGGGCAGCTCTAAGTAATAGCCCTGCTGATGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTCCAAATTCGAATTCATTACTTCCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTCTGCTGCAACGACTAATAAGGCTAGCCAGACTCCGACGACCAACATCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTAAACAGATCTTTTCTTCTCAATATGAAGTCTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAGAGCTTAAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTACCAAGTTTCTCTGCTATGGCTAAACGAGCTCTATTCCTTTGATGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
QY 1102 GTTCCAGGAGTATAGGGGGAGGAGTGTCTGCTGTTCCAGATGGGACGAGGAGTCTCA 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTTCAACAGAAAGATCCAGTAGTAGTAAAGTTTTTCCAGAATACTCGCGTAGAGTTT 1221
Db 261 SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACCTAGCCCGAGTAGGAGGAGGATTTATCTCTACGGAAACGTTGCTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAAAAACCTTGTCTTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320


```
QY 1342 CAACCAACAGTGGACAGCTCTTAATACGAGTAATAATTACGAGATGGAGGAGCTATC 1401
Db |||||:::|||||
QY 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340
Db |||||:::|||||
QY 1402 TTCTGTAAGAATGGTGGCCAA---CCAGGATCCAAATACTCGATCAGTTTCCTTTGAT 1458
Db |||||:::|||||
QY 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
Db |||||:::|||||
QY 1459 GGAGAGGAGTAGTGTTCCTTTAGTAGCAATGATGCTGTGGAAAGGGGAGCTATTTAT 1518
Db |||||:::|||||
QY 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380
Db |||||:::|||||
QY 1519 GCCAAAAGCTCTCGTTGCTAACTGCTGGCCCTGTACAAATTTTAAAGCAATATCGCTAAT 1578
Db |||||:::|||||
QY 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
Db |||||:::|||||
QY 1579 GATGCTGGAGCGCATTTATTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTTATGGA 1638
Db |||||:::|||||
QY 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420
Db |||||:::|||||
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAAATGCTGCCGATGTTAAT 1698
Db |||||:::|||||
QY 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
Db |||||:::|||||
QY 1699 GCGTAACTGTCTCTCAAGCCATTTTCGATGGGATCGGAGGGAATAATACGACATTA 1758
Db |||||:::|||||
QY 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
Db |||||:::|||||
QY 1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCATCGAGATGGCAACCGAAAT 1818
Db |||||:::|||||
QY 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
Db |||||:::|||||
QY 1819 AACCCAGCCAGCCAGCTCTTCCAAACTCTTAAATAATTAACGATGGTGAAGGATACACAGG 1878
Db |||||:::|||||
QY 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
Db |||||:::|||||
QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACGAAGA 1938
Db |||||:::|||||
QY 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Db |||||:::|||||
QY 1939 AGDATGTTCTTCGTGAAAGGCCAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
Db |||||:::|||||
QY 521 ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Db |||||:::|||||
QY 1999 AGTCTGTATATGAAGCTGGAGTACATGGGATTTTGTAACTCCCAACACCACCAACAG 2058
Db |||||:::|||||
QY 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
Db |||||:::|||||
QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCATTTCTCTTCTCTTTG 2118
Db |||||:::|||||
QY 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
Db |||||:::|||||
QY 2119 TTAGCAAAACAATGCAAGTACGAATCTCTTACCAATCTCCAGCGCAGAGATTTCTATCT 2178
Db |||||:::|||||
QY 581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
Db |||||:::|||||
QY 2179 CGAGTCATTGGTAGCACAACTCGTCTGTGTACAAATPAGTGGCCCTATCTTTTGGAG 2238
Db |||||:::|||||
QY 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Db |||||:::|||||
QY 2239 GATTTGGATGATACAGCTTATCATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298
Db |||||:::|||||
QY 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrPheLeuGlySerAsnGlnLysIleAsp 640
Db |||||:::|||||
QY 2299 GTCTCGAAATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCCATCAGATTTGACTCTA 2358
Db |||||:::|||||
QY 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Db |||||:::|||||
QY 2359 GCGAATAGATCCCTAAGTATCGCTATCAAGGAAGCTGGAAGCTGCGTGGGATCCTAAT 2418
Db |||||:::|||||
QY 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
Db |||||:::|||||
QY 2419 ACAGCAATAATGGTCTTATCTCTGAAAGCTACATGGACTAAAACTGGGTATAATCTCT 2478
```

```
Db |||||:::|||||
QY 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
Db |||||:::|||||
QY 2479 GGGCTGAGCGAGTAGCTTCTTTGTTCCAAATAGTTTATGGGATCCATTTTATAGATATA 2538
Db |||||:::|||||
QY 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
Db |||||:::|||||
QY 2539 CGATCTGGCATTTCCAGCAATTCAGCAAGTGTGATGGCGCTCTTATGTGCGAGGATTA 2598
Db |||||:::|||||
QY 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Db |||||:::|||||
QY 2599 TGGGTTTCTGAGTTCGAAATTTCTTCTATCATGACCGCGATGCTTTTAGGTCAAGGATAT 2658
Db |||||:::|||||
QY 741 TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Db |||||:::|||||
QY 2659 CGGTATATTAGTGGGGTTATTCTCTTAGGACAACTCTACTTTGGATCATCGATGTTT 2718
Db |||||:::|||||
QY 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Db |||||:::|||||
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGATGTGTGTCGTTCCAAT 2778
Db |||||:::|||||
QY 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Db |||||:::|||||
QY 2779 CATCATGCTTCATAGGATCCGTTTATCTACTACCCAAAGCTTTTATGTGGATCTCTAT 2838
Db |||||:::|||||
QY 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Db |||||:::|||||
QY 2839 TTGTTCCGAGATGCGTTTATCCGTCTAGCTACGCGTTTGGGAATTCAGCATATGAAAAAC 2898
Db |||||:::|||||
QY 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Db |||||:::|||||
QY 2899 TCATATACATTTGCAGAGGAGGAGGATGTTGTTGGGATATAACTGTCGTGGTGGAGAG 2958
Db |||||:::|||||
QY 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
Db |||||:::|||||
QY 2959 ATTGAGCGGATTTACCGATTTGTGATTACTCCATCTAAGCTCTATTTCAATGAGTTGCGT 3018
Db |||||:::|||||
QY 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Db |||||:::|||||
QY 3019 CTTTTCGTGCAAGCTGAGTTTTTCTTATGCCGATCATGAATCTTTTACAGAGAAAGCGCAT 3078
Db |||||:::|||||
QY 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Db |||||:::|||||
QY 3079 CAAGTCCGGCATTCACAGCGGACATCTCTTAATCTCATGTTCTCTGTTGGAGTGAAG 3138
Db |||||:::|||||
QY 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Db |||||:::|||||
QY 3139 TTTGATCATGTTCTAGTACACATCTCTAATAATATAGCTTTATGGCGCTTATATCTGT 3198
Db |||||:::|||||
QY 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Db |||||:::|||||
QY 3199 GATGCTTATCGACCATCTCTGGTACTGAGACAACGCTCCTATCCATCCACAGAGACATGG 3258
Db |||||:::|||||
QY 941 AspAlaTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960
Db |||||:::|||||
QY 3259 ACAACAGATGCTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGATGCT 3318
Db |||||:::|||||
QY 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Db |||||:::|||||
QY 3319 TCTCTAAACAAGTAATATAGAGATATATGCCATGGAAGATATGAGTATCGATGCTTCT 3378
Db |||||:::|||||
QY 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyAlaTyrGlyLysArgAspThrSer 1000
Db |||||:::|||||
QY 3379 CGAGGCTATGTTTGGAGTCGAGGAAGTAGAGTCGCGTTTC 3417
Db |||||:::|||||
QY 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
Db |||||:::|||||
RESULT 6
ADD43708
ID ADD43708 standard; protein; 1013 AA.
XX
AC ADD43708;
XX
```

DT 15-JAN-2004 (first entry)
 XX Chlamydia trachomatis immunogenic protein, SEQ ID No 3.
 XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
 KW gene therapy; antibacterial.
 XX Chlamydia trachomatis.
 OS
 XX WO2003049762-A2.
 XX
 XX PD 19-JUN-2003.
 XX
 XX PF 12-DEC-2002; 2002WO-1B005761.
 XX
 XX PR 12-DEC-2001; 2001GB-00029732.
 PR 06-AUG-2002; 2002GB-00018233.
 PR 14-AUG-2002; 2002GB-00018924.
 XX
 XX (CHIR-) CHIRON SPA.
 PA
 XX Grandi G, Ratti G;
 XX WPI: 2003-532882/50.
 DR N-PSDB; ADD43709.
 DR
 XX New immunogenic composition having a protein or encoding nucleic acid,
 PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
 PT infection.
 XX
 XX Claim 6; SEQ ID NO 3; 164pp; English.
 PS
 XX The invention relates to a novel immunogenic composition comprising a
 CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
 CC acid comprises any of 131 fully defined amino acid or nucleotide
 CC sequences given in the specification, or has 50% or greater sequence
 CC identity to it, or their fragments. The protein and/or nucleic acid of
 CC the immunogenic composition is useful in the manufacture of a medicament
 CC for the treatment or prevention of infection due to Chlamydia
 CC trachomatis. The infection is treated or prevented by the medicament
 CC elementary body, or for neutralising C. trachomatis elementary bodies,
 CC hence the immunogenic composition can be used in creating a vaccine. The
 CC immunogenic compositions can also be used for the diagnosis of C.
 CC trachomatis infection. The nucleic acids of the immunogenic compositions
 CC can be used to treat disorders by gene therapy. The immunogenic
 CC compositions have antibacterial activity. This sequence represents one of
 CC the 131 C. trachomatis proteins with immunogenic properties of the
 CC invention.
 XX
 SQ Sequence 1013 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5131.50 Matches: 983
 Percent Similarity: 98.8% Conservative: 18
 Best Local Similarity: 97.0% Mismatches: 11
 Query Match: 65.1% Indels: 1
 DB: 7 Gaps: 1
 US-10-701-844-1 (1-4435) x ADD43708 (1-1013)
 QY 382 ATGCACACGCTTTCCATAAGTCTTTCTTCAATGATTCAGCTTATCTTGCTGCTCT 441
 Db 1 MetGlnThrSerPheHisLeuPhePheLeuSerMetileLeuAlaTyrSerCysCysSer 20
 QY 442 TTAATGGGGGGGATATGACGACGATCATGTTCTTCAAGCAATTTACGATGGGAG 501
 Db 21 LeuSerGlyGlyTyrAlaAlaGluileMetileProGlnGlyIleTyrAspGlyGlu 40
 QY 502 ACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTT 561
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60

QY 562 TCTGCAGGAGGTTAAACATTAAATAATCTTGACAAATCTATTGACAGCTTTGCTTTAAGT 621
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
 QY 622 TGTGTTGGAACTTATTAGGGAGTTTACTGTTTATAGGAGGAGGACACTGGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 QY 682 GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAAGTAAATAGCGCTGCTGATGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 QY 742 TTTTACTATTGAGGGTTTAAAGAATTATCTCTTTTCAATTTGCAATTCATTCTGCGGTA 801
 Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
 QY 802 CTGCGCTGTCGAACGACTTAATAGGGTAGCCAGACTCCGACGACCATCTACACGCTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
 QY 862 AATGCTACTATTATTCTAAACACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTC 921
 Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180
 QY 922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAAGAGCTTAACGGTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
 QY 982 GGAATTTAGCAAGCTTTGCTCTTCCAGAAATATCTGCTCAAGCTGATGGGAGCTTGT 1041
 Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
 QY 1042 CAAGTAGTCACCAAGTTTCTCTGCTAGCTTAACGAGGCTCTATTGCTTTGAGCGAAT 1101
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProlleAlaPheIleAlaAsn 240
 QY 1102 GTTGCAGGAGTAGTAAGCGGGAGGATTGCTGCTGTTTCCAGGATGGGACGAGGAGTGTCA 1161
 Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
 QY 1162 TCATCTACTTCAACAGAAATCCAGTAGTAAGTTTTCAGAAATATCTGCGGTAGAGTTT 1221
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 QY 1222 GATGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAGAGCTTGTCTTCCTG 1281
 Db 281 AspGlyAsnValAlaAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
 QY 1282 AATAATGGAAAAACCTTTGTTCTCAACAATGTTGCTTCTCTCTTACATTGCTGCTAAG 1341
 Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320
 QY 1342 CAACCAACAAAGTGGACAGGCTTCTAATACAGTAATAATTAACGAGATGAGGAGCTATC 1401
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
 QY 1402 TTCTGTAAGAACTGGTGGCGAA---GCAGGATCCAATAACTCTGATCAGTTTCTCTTGTAT 1458
 Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 QY 1459 GGAGAGGAGCTAGTTTCTTTTAGTACCAATGTTAGCTCTGGAAAGGGGAGCTATTATAT 1518
 Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
 QY 1519 GCCAAAAGCTCTCGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578
 Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 QY 1579 GATGGTGGAGCGATTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420

QY 1639 GATATTATTTTCGATGGGAATCTTAAAGACAGCAAGAGAAATGCTGCCGATCTTAAT 1698
DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAspValAsn 440
QY 1699 GCGTAACCTGTGCTCCAGAGCCATTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyValIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACGAGCCAGCCAGCTCTCCAACTTCTAATAATTAACGATGGTGAAGGATACACAGG 1878
DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyThrGly 500
QY 1879 GATATTGTTTTCGTAAGGAGCAGTACTTGTATCCAAATGTTTACGATAGACGAAGA 1938
DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuThrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTGTTTTCGTAAGGCAAAATATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998
DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGATATGAGAGCTGGGAGTACATGGGATTTGTAACTCCACACCAACCAACAG 2058
DB 541 SerLeuThrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CTTCTGCGCGTAATCAGTTGATCAGCTTTCGAATCTGCATTTGTCTCTTCTCTTTG 2118
DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAAAACAATGCAAGTACGAATCTCTACAAATCTCCAGCGCAAGATTCATCT 2178
DB 581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 CAGTCATTGTTAGCAACTCTGCTGTTCTGTTACAAATAGTGGGCTATCTTTTGG 2238
DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGTGATGATACAGCTTATCATAGTATGATGGTAGTTCTTAATCAAAAATCAAT 2298
DB 621 AspLeuAspThrAlaThrAspArgGlyAspTrpLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTGAAAATACAGTTAGGACTAGACCCCGAGCTAATGCCCCATCAGATTGACTCTA 2358
DB 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGATGAGATGCTTAAGTATGGCTATCAAGNAGCTGCAAGCTGCGTGGGATCCTAAT 2418
DB 661 GlyAsnGluMetProLysThrGlyThrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAAAATAATGCTCTTATCTCTGAAAGCTACATGCACTAAAGCTGGTATATCTCT 2478
DB 681 ThrAlaAsnAsnGlyProThrThrLeuLysAlaThrTrpThrLysThrGlyThrAsnPro 700
QY 2479 GGGCTGAGCAGTAGCTCTTTGTTTCCAAATAGTTTATGGGATGCCATTTTAGATATA 2538
DB 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGCCCATTCAGCAATTCAGCAAGTGTGGATGGGCGCTCTTATTGTGAGGATTA 2598
DB 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerThrCysArgGlyLeu 740
QY 2599 TGGGTTCTCTGGAGTTTCGAATTTCTCTATCATGACCGCATGCTTTTAGTCAGGATAT 2658
DB 741 TrpValSerGlyValSerAsnPheThrHisAspArgAspAlaLeuGlnGlyThr 760
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTCTAATTTGGATCATCGATTTT 2718
DB 761 ArgThrIleSerGlyThrSerLeuGlyAlaAsnSerThrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTGTTAGTATCTAAAGATTTATGTAGTGTGCTCCAAAT 2778

DB 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspThrValValCysArgSerAsn 800
QY 2779 CATCATGCTTCATAGGATCGGTTTATCTATCTACCAACAAAGCTTTATGCGATCTTAT 2838
DB 801 HisHisAlaCysIleGlySerValThrLeuSerThrLysGlnAlaLeuCysGlySerThr 820
QY 2839 TTGTTGCGAGATCGGTTTATCGGTTCTAGCTACCGGTTTGGGATTCACCATATGAAACC 2898
DB 821 LeuPheGlyAspAlaPheIleArgAlaSerThrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGACAGAGAGAGGATGTTCTGTTGGGATAATAACTGTCCTGCTGAGAG 2958
DB 841 SerThrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGATACCGATTTGATTAATCTCAATCTCAAGCTCTAATTTGAATGAGTTGGT 3018
DB 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuThrLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTCGAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT 3078
DB 881 ProPheValGlnAlaGluPheSerThrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCCTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138
DB 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCATGTTCTTACATCATCTTAATAATATAGCTTTATGCGGCTTATATCTGT 3198
DB 921 PheAspArgCysSerSerThrHisProAsnLysThrSerPheMetGlyAlaThrIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAAGCTCCTATCCATCCAGAGACATGG 3258
DB 941 AspAlaThrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTCTTTCATTAGCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCT 3318
DB 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetThrAla 980
QY 3319 TCTCTAACAGTAATATAGAGTATATGCGCAATGGAAGATATGATATCGAGATGCTTCT 3378
DB 981 SerLeuThrSerAsnIleGluValThrGlyHisGlyArgThrGluThrArgAspThrSer 1000
QY 3379 CGAGCTATGTTTTCAGTGCAGGAGTAGATGCTCCGGTTC 3417
DB 1001 ArgGlyThrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 7
ADD43714
ID ADD43714 standard; protein; 1013 AA.
XX AC ADD43714;
XX DT 15-JAN-2004 (first entry)
XX DE Chlamydia trachomatis immunogenic protein, SEQ ID No 9.
XX XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
XX KW gene therapy; antibacterial.
XX OS Chlamydia trachomatis.
XX PN WO2003049762-A2.
XX PD 19-JUN-2003.
XX PF 12-DEC-2002; 2002WO-IB005761.
XX XX 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX XX (CHIR-) CHIRON SPA.

XX Grandi G, Ratti G;
 XX WPI; 2003-532882/50.
 DR N-PSDB; ADD43715.
 XX New immunogenic composition having a protein or encoding nucleic acid,
 PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
 PT infection.
 XX Claim 6; SEQ ID NO 9; 164pp; English.
 XX The invention relates to a novel immunogenic composition comprising a
 CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
 CC acid comprises any of 131 fully defined amino acid or nucleotide
 CC sequences given in the specification, or has 50% or greater sequence
 CC identity to it, or their fragments. The protein and/or nucleic acid of
 CC the immunogenic composition is useful in the manufacture of a medicament
 CC for the treatment or prevention of infection due to Chlamydia
 CC trachomatis. The infection is treated or prevented by the medicament
 CC eliciting an immune response which is specific to a C. trachomatis
 CC elementary body, or for neutralising C. trachomatis elementary bodies,
 CC hence the immunogenic composition can be used in creating a vaccine. The
 CC immunogenic compositions can also be used for the diagnosis of C.
 CC trachomatis infection. The nucleic acids of the immunogenic compositions
 CC can be used to treat disorders by gene therapy. The immunogenic
 CC compositions have antibacterial activity. This sequence represents one of
 CC the 131 C. trachomatis proteins with immunogenic properties of the
 XX invention.
 XX Sequence 1013 AA;
 SQ

Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5131.50 Matches: 983
 Percent Similarity: 98.8% Conservative: 18
 Best Local Similarity: 97.0% Mismatches: 11
 Query Match: 65.1% Indels: 1
 DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD43714 (1-1013)

QY 382 ATGCAAACTCTTCCATAGTTCTTTCTTCAATGATTCTAGCTTATTTCTGCTCT 441
 Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaIyrSerCysSer 20
 QY 442 TTAATGGGGGGATATGACAGAAATCATGGTTCCTCAAGAAATTTACGATGGGGAG 501
 Db 21 LeuSerGlyGlyTyralaAlaGluIleMetIleProGlnGlyIleTyraPheGlyGlu 40
 QY 502 AGTTAACTGTATCTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTT 561
 Db 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
 QY 562 TCTGACGGAGGTACATTAATAATCTTGCAATCTTATTCAGCTTTCCTTTAAGT 621
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
 QY 622 TGTTTTGGGAACCTATTAGGGAGTTTACTGTTTATAGGAGAGACACTCGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 QY 682 GAGAACATACGGACTTCTACAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGAGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 QY 742 TTTACTATTGAGGGTTTAAAGAAATTCCTTTTCCAAATTCGAATTCATTTACCTGCGTA 801
 Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
 QY 802 CTCCTGCTCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAATCTACACCGTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160

QY 862 AATGCTACTATTTTCTTAAACAGATCTTTTGTGTACTCAATAATGAGAAGTTCATTC 921
 Db 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
 QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
 QY 982 GGAATTAGCAAGCTTTGTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
 Db 201 GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
 QY 1042 CAAGTAGTCACCAAGTTCTCTGTATGCTAACGAGGCTCTTATTCCTTTCTAGCGAAT 1101
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
 QY 1102 GTTCAGAGTAAGAGGGGGAGTTCTGCTGTTTCAGGATGGGCACGACGAGGTGCA 1161
 Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
 QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAACTTTTCCAGAAATACTGCGGTAGAGTTT 1221
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 QY 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTTACTCTACGGGAACGTCTCTTCCTG 1281
 Db 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300
 QY 1282 AATAATGAAAAACCTGTTTCTCAACAAATGTTGCTTCTCTGTTTACATTCGCTGTAAG 1341
 Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyIleAlaAlaGlu 320
 QY 1342 CAACCAACAGTGGACAGCTTCTAATAGAGTAATAATTACGGAGATGAGGAGCTATC 1401
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle 340
 QY 1402 TTCTCTAAGAATGGTGGCGCAA---GCAGGATCCAATACTCTGGATCAGTTTCTCTTGTAT 1458
 Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 QY 1459 GGAGGAGGAGTAGTTTCTTTAGTAGCAATAGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
 Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTy 380
 QY 1519 GCCAAAAGCTCTCGGTTGCTTAACGTGCGCTGTACATTTTAAAGGAATATCGCTAAT 1578
 Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 QY 1579 GATGTTGGAGCGCATTTATTAGGAGAACTCTGAGAGACTCAGTTTATCTCTGATTATGGA 1638
 Db 401 AspGlyGlyAlaIleTyLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyGly 420
 QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCAAGAGAGATGCTCCGAGTGTAAAT 1698
 Db 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
 QY 1699 GCGTAACTGTGCTCACAACCCATTTTCGATGGATCCGGAGGGGAAATAACGACATTA 1758
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIleThrThrLeu 460
 QY 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAAT 1818
 Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
 QY 1819 AACCCAGCAGCGCAGTCTTCCAAACTTCTAAATAATTAACGATGGTGAAGGATACACAGG 1878
 Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly 500
 QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTCTGACCAAAATGTTACGATAGGACGAGA 1938
 Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly 520

QY 1939 AGGATTGTTCTTCGTAAGAGCAAAATATCAGTGAATTCCTCAAGTCAGACAGTGGG 1998
DB |||||
521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCGTATATCGAAGCTGGAGTACATCGGATTTTGTAACTCCCAACACCAACAG 2058
DB |||||
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTGCGCGTAATCAGTTGATCAAGCTTTCCTCAATCTGCATTTCTCTTCTTTG 2118
DB |||||
561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAACAATGCAAGTACGAATCCTCTACCAATCCTCCAGCCGACAGATTCATCCT 2178
DB |||||
581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProThrAsnProAlaGlnAspSerHisPro 600
QY 2179 CGAGTCATTGGTAGCAACAACCTCTCGTTCTGTGTACAAATAGTGGGCCCTATCTTTTTCAG 2238
DB |||||
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATPACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298
DB |||||
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTCTGAAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTGACTCTA 2358
DB |||||
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGAATAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTGGCGGATCCTAAT 2418
DB |||||
661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAAAATATGGTCTTATCTCTGAAAGCTACATGAGCTAAACCTGGGTATATCCT 2478
DB |||||
681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCTGAGCAGTAGTCTTCTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
DB |||||
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CQATCTGCGCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTTGTCCAGGATTA 2598
DB |||||
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTCTGGAGTTTCGAATTTCTTCTATCAGACCCGATGCTTTAGGTCAAGGATAT 2658
DB |||||
741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCCTACTTTGGATCATCGATGTTT 2718
DB |||||
761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGTAGATCTTAAAGATTATGATGTGTCGTTCCCAAT 2778
DB |||||
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTGTCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGATCCTAT 2838
DB |||||
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGGAGATCGGTTTATCCGTCTAGCTAGCGGTTTGGGAATCAGCATATGAACCC 2898
DB |||||
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGAGCGATGTTCTGTGGATAATAACTGCTGCTGCGAGAG 2958
DB |||||
841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTACCGATTGTGATTACTCCATCTTAAGCTCTATTGATGAGTTGCGT 3018
DB |||||
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGCAT 3078

DB |||||
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCAITTCAGAGCGGCACATCTCCCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138
DB |||||
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTGATCGATGTTCTAGTACATCCCTTAATAATAATAGCTTTATGGCGCTTATATCTGT 3198
DB |||||
921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCCACCATCTCTGGTACTGAGACAAAGCTCTCATCCATCAAGAGACATGG 3258
DB |||||
941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTTAGCAAGACATCGAGTGTGGTAGAGGATCTATGATGCT 3318
DB |||||
961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTATATAGAAGTATATGGCCATCGAAGATATGAGTATCGAGATGCTTCT 3378
DB |||||
981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGTTTGGTGCAGGAAGTAGAGTCCGGTTC 3417
DB |||||
1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 8
ADD43786
ID ADD43786 standard; protein; 1013 AA.
XX
AC ADD43786;
XX
DT 15-JAN-2004 (first entry)
XX
DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 81.
XX
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO2003049762-A2.
XX
PD 19-JUN-2003.
XX
PF 12-DEC-2002; 2002WO-IB005761.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Ratti G;
XX
DR WPI; 2003-532882/50.
DR N-PSDB; ADD43787.
XX
PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
PS Claim 6; SEQ ID NO 81; 164pp; English.
XX
CC The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament

CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic composition can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of
CC the 131 C. trachomatis proteins with immunogenic properties of the
CC invention.
XX

SQ Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD43786 (1-1013)

QY 382 ATGCAACGCTCTTCCATAGTTCTTTCTTCAATGATCTAGCTTATCTTCTGCTCT 441
DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATGGGGGGATATGACGAGAAATCATGGTCTCCAGGAATTTACGATGGGAG 501
DB 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 AGCTTAACGTATCTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrValPhe 60
QY 562 TCTGACGAGAGTTAACTTAAAAATCTTGCAATTTCTATTCAGCTTTCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
QY 622 TGTGTTGGAACTTATAGGAGTTTACTGTTTTAGGGAGGAGCACTGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACTACGCTCTCAATAGGGAGCTCAAGTAATAGCGTCTGCTGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTACTTGCCTGA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCyAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTAATAAGGGTAGCAGACTCCGACGACCAATCTACACGCTCT 861
DB 141 LeuProAlaAlaThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
QY 862 AATGGTACTATTATCTTAAACAGATCTTTGTTACTCAATATGAGAAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGACTTTAAACGTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGCTCTTCCAAAGAAATACTGCTCAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTACCAGTTTCTCTGCTATGCTGCTAACGAGCTCTTATGCTTGTGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
QY 1102 GTTGACGAGTAAAGAGGGAGGATTTGCTGCTTTCAGGATGGGACGAGGAGTGTCA 1161
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlyValSer 260

QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCACGAAATATCTGCGGTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTACTCTACGGAGAGTGTCTCTCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATATGCAAAACCTCTGTTCTCAACAATGTTGCTCTCTCTGTTTACATGCTCTAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
QY 1342 CAACCAACAAGTGGACAGGCTTCTAATACAGTAAATAATTACGGAGATGGAGGAGTATC 1401
DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTCTCTAAGATGTTGGCGCAA---GCAGATCCAATACTCTGGATCAGTTTCTCTTGTAT 1458
DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 CGAGAGGAGTAGTTTCTTTAGTAGCAATGAGCTGCTGGCAAGGGGAGCTATTTAT 1518
DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380
QY 1519 GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAATTTTAAAGAAATATCGCTAAT 1578
DB 381 AlaLysLysLeuSerValAlaAsnCySGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATGTGGAGCCATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTCTGATTTATGCA 1638
DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTTATTTTCGATGGGAATCTTAAAGAAACAGCAAGAGAGATCTGCGCATGTAAAT 1698
DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGCTAACTGTCTCTCAACAGCCATTTTCGATGGGATCGGGAGGGAATAAATACGACATTA 1758
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAAGCGGGGATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACCCAGCGGCGAGCTTCCAACTTCTAAATAATTAACGATGTCGAGGATACACAGGG 1878
DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500
QY 1879 GATATTTGTTTTCGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGGA 1938
DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTTGTTCTTCGTAAGGCAAAATATATAGTGAATTTCTCTAAGTCAGACAGGTTGG 1998
DB 521 ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGAAGCTGGGAGTACATGGATTTTGTAACTCCACAAACCAACACACAG 2058
DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTCAATCTGCAATTTGTCTCTTTCTTCTTTG 2118
DB 561 ProProAlaAlaAsnGlnIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTACGAAACAATCGAGTTACGAATCTCTTACCAATCTCTCAGCGCAAGATTCTCATCT 2178
DB 581 LeuAlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCGATCTGTTGTTAGCACACTGCTGTTCTGTTTACAAATAGTGGCTATCTTTTGTAG 2238
DB 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProfilePheGlu 620


```
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAATCAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCCCTGAATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGATAGATGCCCTAAGTATGCTATCAAGGAAGCTGGAGCTGGCGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAATATAGTCTTATCTGAAAGCTACATGACTAAACCTGGGTATATCTCT 2478
Db 681 ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnPro 700
QY 2479 GGGCTGAGCAGTAGCTTCTTTGGTCCAAATAGTTTATGGGATCCATTTTACATATA 2538
Db 701 GlyProGluAaGValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCAATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGATT 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740
QY 2599 TGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCGGATGCTTTAGTACGGGAT 2658
Db 741 TrpValSerGlyValSerAsnPheTyThrHisAspArgAspAlaLeuGlyGlnGlyTy 760
QY 2659 CGGTATATTAGTGGGGTTTATCTCTTAGGAGCAAACTCTACTTTGGATCATCGATGTT 2718
Db 761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAsnSerTyPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTATGTAGTGTGTCGTTCCA 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsn 800
QY 2779 CATCATGCTTGATAGGATCCGTTTATCTATCTATCTACCAAGCTTTATGCGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820
QY 2839 TTGTTCCGAGATGCGTTTATCCGTCTAGCTAGCGGTTTGGCAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGAGCGATGTTCTGCGGATTAATACTGTCTGCGTGGAG 2958
Db 841 SerTyThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTACCGATTGTGATTACTCATCTAAGCTCTATTTGAATGAGTTGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGATCTTTTACAGAGAACGCGAT 3078
Db 881 PropheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGGCGCTTATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTySerPheMetGlyAlaTyIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGCTACTGACAGACGCTCTATCCCATCAAGAGACATGG 3258
Db 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCGCTTTTCATTTAGCAGACATGGAGTTGTGGTTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980
QY 3319 TCTCTAACAGTAAATATAGATATATGCGCATGGAAGATATGAGTATCAGATGCTTCT 3378
```

```
Db 981 SerLeuThrSerAsnIleGluValTyGlyHisGlyArgTyGlyTyArgAspThrSer 1000
QY 3379 CGAGCTATGTTTGGAGTCGAGGAAGTAGAGTCGGTTC 3417
Db 1001 ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 9
ADD43788
ID ADD43788 standard; protein; 1013 AA.
XX
AC ADD43788;
XX
DT 15-JAN-2004 (first entry)
DE Chlamydia trachomatis immunogenic protein, SEQ ID No 83.
XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
XX Chlamydia trachomatis.
OS
PN WO2003049762-A2.
XX
PD 19-JUN-2003.
XX
PF 12-DEC-2002; 2002WO-IB005761.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Ratti G;
XX
WPI; 2003-532882/50.
DR N-PSDB; ADD43789.
XX
New immunogenic composition having a protein or encoding nucleic acid,
useful for diagnosing, preventing and/or treating Chlamydia trachomatis
infection.
PS Claim 6; SEQ ID NO 83; 164pp; English.
XX
The invention relates to a novel immunogenic composition comprising a
protein or nucleic acid, and an adjuvant, where the protein or nucleic
acid comprises any of 131 fully defined amino acid or nucleotide
sequences given in the specification, or has 50% or greater sequence
identity to it, or their fragments. The protein and/or nucleic acid of
the immunogenic composition is useful in the manufacture of a medicament
for the treatment or prevention of infection due to Chlamydia
trachomatis. The infection is treated or prevented by the medicament
eliciting an immune response which is specific to a C. trachomatis
elementary body, or for neutralising C. trachomatis elementary bodies,
hence the immunogenic composition can be used in creating a vaccine. The
immunogenic compositions can also be used for the diagnosis of C.
trachomatis infection. The nucleic acids of the immunogenic compositions
can be used to treat disorders by gene therapy. The immunogenic
compositions have antibacterial activity. This sequence represents one of
the 131 C. trachomatis proteins with immunogenic properties of the
invention.
SQ Sequence 1013 AA;
XX
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1
```

US-10-701-844-1 (1-4435) x ADD43788 (1-1013)

382 ATGCAAACTCTTTCATAGTTCTTTCTTCAATGATCTACGTTATTTCTTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
442 TTAAATGGGGGGATATGACAGAGAAATCATGGTTCCTCAAGAAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
502 ACGTTAACTGATCATTTCCCTATACCTGTTATAGGAGATCCGAGTGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
562 TCTGCAGGAGAGTTAATCAATTAATAATCTTGACAAATCTATTCACGTTTGGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
622 TGTGTTTGGGAACCTTATTAGGAGCTTTTACTGTTTATAGGAGAGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
682 GAGAACATACGCACTCTCAAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGCACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
742 TTTACTATTAGGGTTTTAAGAATTATTCCTTTTCCAAATGCAATTCATTACTTGGCTTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnGlyAsnSerLeuLeuAlaVal 140
802 CTGCGCTGTCACAGCACTAATAGGCTAGCCAGACTCCGACGACACATCTACACGCTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
862 AATGGTACTATTATTTCAAAACAGATCTTTTGTACTCAATATAGAAAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
922 TATAGTAATTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGAGCTTAAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaAlaIleAspAlaLysSerLeuThrValGln 200
982 GGAATTAGCAAGCTTTGTCTCCAGAAATACTCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
1042 CAAGTAGTCACCACTTCTCTGCTATGCTTAACGAGGCTCTATTGCTCTTTGAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
1102 GTTGCAGGAGTAAGAGGGGAGGATGCTGCTGTTCCAGATGGGACGAGGAGTGTCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACCTGCGTAGACTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
1222 GATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTACGGAACGTTGCTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
1282 AATAATGGAAAACTTGTCTTCTCAACAATGTGCTTCTCTGTTTATCATGCTGCTGAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320
1342 CAACCAACAGTGACAGGCTCTTAATACGAGTAATTAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
1402 TTCTGTGAAGATGCTGGCGCAA--GCAGATCCAAATCACTGATCACTGATTTCTTTGAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360

1459 GGAGAGGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyGlyGlyAlaIleTyr 380
1519 GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTATCAATTTTAAAGGAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
1579 GATGTGGAGCCATTATTATTAGAGAACTCTGAGAGAGCTCAGTTTATCTCTGATATGCA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLysSerAlaAspTyrGly 420
1639 GATATTATTTCGATGGGAATCTTAAAGAACAGCAACAGCAAGAGAGATGCTCCGATGTTAAT 1698
Db 421 AspllellePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
1699 GCGCTAACTGTCTCTCAACAGCCATTTCGATGGGATCGGAGGAGGAGGAGGAGGAGGAG 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTTAATGATCCCATCGAGATGCGAAACGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
1819 AACCAGCCAGCGAGCTTCCAAACTTCTAAAAATTAACGATGCTGAAGGATACACAGGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
1879 GATATTGTTTTGCTTAATGGAAGCAGTACTTTGTACAAAATGTTACGATAGAGCAAGGA 1938
Db 501 AspllellePheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
1939 AGGATTGTTCTGCTGAAAGCGCAAAATTAATCATGTCGAATCTCTAAGTCAGAGGAGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCAACCAACCAACCAACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCTCAATCTGCAATTCCTCTCTCTCTCTT 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
2119 TTAGCAAAACATCGAGTTACGAATCTCTCTACCAATCTCCAGGCAAGATTTCTCATCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
2179 GCAGTCAATGGTAGCACTGCTGTTCTGTACAAATAGTGGGCTATCTTTTGTGAG 2238
Db 601 AlaIlelleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
2239 GATTGATCATACAGCTTATGATAGTATGATGCTAGTGTCTAGTGTCTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
2299 GTCTCTGAAATTAAGTTAGGAGCTAAGCCCGGAGCTTAATCCCATCATGATGCTACTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAlaAsnAlaProSerAspLeuThrLeu 660
2359 GGGATGAGATGCTTAAGTATGCTATCAAGGAGCTGGAAGCTTGCTGGATCTCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
2419 ACAGCAAAATAGTCTCTTATCTCTGAAAAGCTACATGAGCTAAAACTGGGTATAAATCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
2479 GGGCTCTGAGCGAGTAGTCTTTGGTTCGAAATAGTTATGGGATCCATTTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720

```
QY 2539 CGATCTGGCCATTTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGATT 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTAGGTCAGGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGAGCAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAATTTGGTAGACTAAAGATTATGTAGTGTGCTTCCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATACCAACCAAGCTTTATGTGGATCCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGAGATGCGTTTATCCGTCCTAGCTACGAGTTCGGGTTTGGAAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGAGCGATGTTCTGGGATAATAACTGTCTGGCTCGAGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTACCGATTCGATCTACTCCATCTTAAGCTCTATTGAATGAGTTGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CCTTTTCGTCAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCCTAAATCTATCATGTTCTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAAGCTCTATCCCATCAAGAGCATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGCAGACATGAGTGTGGTTAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAATATAGATATATGCGCATATGCGCATATGAGTATGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGTTTGTAGTCAGAGAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 10
ADD43710
ID ADD43710 standard; protein; 1013 AA.
XX AC ADD43710;
XX DT 15-JAN-2004 (first entry)
XX DE Chlamydia trachomatis immunogenic protein, SEQ ID No 5.
XX KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
XX gene therapy; antibacterial.
```

```
OS Chlamydia trachomatis.
XX WO2003049762-A2.
XX 19-JUN-2003.
XX 12-DEC-2002; 2002WO-IB005761.
XX 12-DEC-2001; 2001GB-00029732.
XX 06-AUG-2002; 2002GB-00018233.
XX 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
XX Grandi G, Ratti G;
XX WPI; 2003-532882/50.
XX N-PSDB; ADD43711.
XX New immunogenic composition having a protein or encoding nucleic acid,
XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis
XX infection.
XX Claim 6; SEQ ID NO 5; 164pp; English.
XX The invention relates to a novel immunogenic composition comprising a
XX protein or nucleic acid, and an adjuvant, where the protein or nucleic
XX acid comprises any of 131 fully defined amino acid or nucleotide
XX sequences given in the specification, or has 50% or greater sequence
XX identity to it, or their fragments. The protein and/or nucleic acid of
XX the immunogenic composition is useful in the manufacture of a medicament
XX for the treatment or prevention of infection due to Chlamydia
XX trachomatis. The infection is treated or prevented by the medicament
XX eliciting an immune response which is specific to a C. trachomatis
XX elementary body, or for neutralising C. trachomatis elementary bodies,
XX hence the immunogenic composition can be used in creating a vaccine. The
XX immunogenic compositions can also be used for the diagnosis of C.
XX trachomatis infection. The nucleic acids of the immunogenic compositions
XX can be used to treat disorders by gene therapy. The immunogenic
XX compositions have antibacterial activity. This sequence represents one of
XX the 131 C. trachomatis proteins with immunogenic properties of the
XX invention.
XX SQ Sequence 1013 AA;
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1
US-10-701-844-1 (1-4435) x ADD43710 (1-1013)
QY 382 ATGCAAAAGCTCTTCCATAAGTCTTCTTCAATGATCTAGCTTATTCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGATATGACAGCAAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGTATCATTTCCCTATATCTGTATAGGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGAGGAGAGTTAAACATTAATAAAATCTTGACAAATCTTATTCGAGCTTTCGCTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TCTTTTGGGAACCTTATTAGGGAGTTTACTGTGTTTGGGAGGAGCACCTCGTTGACTTTC 681
```

81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGGACTTCTCAAAATGGGCGAGCTTAAGTAATAGCGCTGCTGATGGACTG 741
Db 101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
QY 742 TTTACTATTGAGGGTTTAAAGAATTATCTTTTCCAAATGCAATTCATTACTTGGCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTTAATAAGGTAGCCAGACTCCGACGACAACTACACCGTCT 861
Db 141 LeuProAlaIaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGGTACTATTTATCTAAACAGATCTTTGTGTACTCAATATGAGAAGTCTTCATTC 921
Db 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAenLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTGCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTCACAGTTCTCTGCTATGGCTAACGAGGCTCTATTTGCCCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240
QY 1102 GTTCAGGAGTAAGGGGGAGGATCTGCTGTTTACAGGATGGCAGCAGGAGTGTC 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATFACCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
QY 1222 GATGGGAAGTACGCCGAGTAGGAGGAGGATTTACTCTCTAGCGGAAGCTTCTTCTG 1281
Db 281 AspGlyAenValAlaArgValGlyGlyIleTySerTySerTyGlyAenValAlaPheLeu 300
QY 1282 AATAATGAAACCTTTCTCAACAATGTGCTCTCTCTGTTTACATTTGCTGCTAAG 1341
Db 301 AsnAenGlyIleThrLeuPheLeuAenAsnValAlaSerProValTyIleAlaAlaGlu 320
QY 1342 CAACCAACAACTGGACAGCTTCTAATACGAGTATAATACGAGATCGGAGAGCTATC 1401
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAenTyGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTAAAGATGGTGGCAA---GCAGATCCAAATACTCTGGATCAGTTTCTTTGAT 1459
Db 341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAsp 360
QY 1459 GGAGGAGGAGTAGTTTCTTTAGTAGCAATGCTAGCTGCTGGAAAGGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyAlaIleTy 380
QY 1519 GCCAAAAAGCTCTCGGTTCCTAACCTGTCGCCCTGTACAAATTTTAAAGAAATATCCCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
QY 1579 GATGCTGGAGCAATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638
Db 401 AspGlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyGly 420
QY 1639 GATATTTATTTGATGGGATCTTAAAGAACAGCAGCAAGAGACTGCTGCCGATGTTAAT 1698
Db 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440
QY 1699 GCGCTAAGTCTGCTCTCAAGCCATTTTCGATGGGATCGGAGGAGAAATTAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460

QY 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAspProIleGluMetAlaAenGlyAen 480
QY 1819 AACCAGCCAGCGAGTCTTCCAAACTTCTAAAAATTAACAGATGCTGAAGATACACAGGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGlyTyThrGly 500
QY 1879 GATATCTTTTTCGTAATGAGCAGTACTTTGTACCAAAATGTTACGATAGACCAAGGA 1938
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyGlnAenValThrIleGlnGlnGly 520
QY 1939 AGGATCTTCTTCTGTAAGAGGCAAAATTAATCAGTGAATCTCTAAGTCAGACAGGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAAACCACCAACAG 2058
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCTCAATCTGCAATTTGCTCTTCTTCTTTG 2118
Db 561 ProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerLeu 580
QY 2119 TTACCAACAAATCAGTTCACGAATCTCTCAACCAATCTCCAGCGCAAGATTTCTCATCT 2178
Db 581 LeuAlaAenAlaValThrAenProThrAenProProAlaGlnAenSerHisPro 600
QY 2179 GCAGTCAATGCTAGCACAACTGCTGTTCTGTTTCAATTAAGTGGGCTTCTTTTTCAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTTGATGATACAGCTTATGATAGTATGATGGCTAGGCTTCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAenGlnLysIleAsp 640
QY 2299 GTCTCTGAAATPACAGTTAGGGACTAAGCCCCAGCTAATGCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660
QY 2359 GGGAAATGAGATGCTCAAGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGATCTTAAT 2418
Db 661 GlyAenGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY 2419 ACAGCAATAATGCTCTTATCTGTAAGCTACATGCACTGCACTAAACCTGGGTATATCTCT 2478
Db 681 ThrAlaAenAenGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnPro 700
QY 2479 GGGCTGAGCGAGTAGCTTCTTGTTCCTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGCGCAATTCAGCAATTAAGCAAGTGTGGATGGGCGCTCTTATTTGTCAGGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740
QY 2599 TGGGTTCTGGAGTTTCAATTTCTTATCATGACCGCGATGCTTTTAGGTACAGGATAT 2658
Db 741 TrpValSerGlyValSerAenPheTyHisAspArgAspAlaLeuGlyGlnGlyTy 760
QY 2659 CGGTATATTTAGTGGGGTTATTTCTTAGGAGCAACTCTCTTTTGGATCATCGATGTTT 2718
Db 761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAenSerTyPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACGAAGTATTTGGTAGATCTAAAGATTTATGTAGTGTGCTGTTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAen 800
QY 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCAACAGCTTTTATGTGGATCTCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820

```
QY 2839 TTGTCGGAGATCGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAAC 2898
DB 821 LeuPheGlyAspAlaPheIleAgaIaSerTyGlyPheGlyAsnGlnHisMetIleThr 840
QY 2899 TCATATACATTTTCAGAGGAGAGCGATGTTCTGGGATATAACTGTCTGGCTGGAGAG 2958
DB 841 SerTyThrPheAlaGluGluSerAspValArgTIPAspAenAenCysLeuValGlyGlu 860
QY 2959 ATTGGAGCGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGATTGGCT 3018
DB 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyLeuAenGluLeuArg 880
QY 3019 CCTTCCTGCAAGCTGATTTTCTTATGCCGATCATGAATCTTTTACAGAGGAGGCGAT 3078
DB 881 ProPheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCGATTCAAGAGCGGACATCTCTAAATCTATCAGTCTCTGTGGAGTGAAG 3138
DB 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyVallys 920
QY 3139 TTGTATCGATGTTCTAGTACATCCTTAATAATATATAGCTTTATGGCGCTTATCTGT 3198
DB 921 PheAspArgCysSerSerThrHisProAsnLysTySerPheMetGlyAlaTyIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGCTGCTGAGACACGCTCTCTATCCCATCAAGACATGG 3258
DB 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTIP 960
QY 3259 ACAACAGATGCTTTCATTTTATGACAGACATGGAGTGTGGTTAGAGATCTATATGCT 3318
DB 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980
QY 3319 TCTCTACAGATATATAGATATATGCGCATATGCGCATATGATATCGAGATGCTTCT 3378
DB 981 SerLeuThrSerAsnIleGluValTyArgIleGlyArgTyGluTyArgAspThrSer 1000
QY 3379 CGAGCGTATGTTGTAGTGACAGGAAGTAGAGTCCGGTTC 3417
DB 1001 ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 11
ADD43706
ID ADD43706 standard; protein; 1013 AA.
XX
AC ADD43706;
XX
DT 15-JAN-2004 (first entry)
XX
DE Chlamydia trachomatis immunogenic protein, SEQ ID No 1.
XX
KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO2003049762-A2.
PD
PP 19-JUN-2003.
XX
PF 12-DEC-2002; 2002WO-IB005761.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Ratti G;
XX
DR WPI; 2003-532882/50.
DR N-P8DB; ADD43707.
XX
PT New immunogenic composition having a protein or encoding nucleic acid,
```

```
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
XX Claim 6; SEQ ID NO 1; 164pp; English.
XX
CC The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of
CC the 131 C. trachomatis proteins with immunogenic properties of the
CC invention.
XX
SQ Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD43706 (1-1013)
QY 382 ATGCAACGCTTTCATATAGTCTTCTTCAATGATCTAGCTATCTTGTGCTCT 441
DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTySerCysCys 20
QY 442 TTAATCGGGGGATATGACAGAAATCATGTTCTCTCAAGAAATTTACAGTGGGAG 501
DB 21 LeuSerGlyGlyTyAlaAlaGluIleMetIleProGlnGlyIleTyAspGlyGlu 40
QY 502 ACGTTAACTGATCAATTTCCCTATCTTATAGAGATCCGAGTGGAGTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGAGAGAGATTAACTTAAAAATCTTGACAAATCTTATGAGCTTTCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTCGGAACTTATAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGACATACCGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGTCTGTGATGAGCTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTACTATGAGGTTTAAAGAAATATCTTTTCAATTCGCAATTCATCTACTCCGCTA 801
DB 121 PheThrIleGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTAAATAAGGGTAGCAGACTCCGACGACACACATCTACACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGCTACTATTATTCTAAACAGATCTTTTGTGTACTCAATATATAGAGAGTTCTCATTC 921
DB 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuLysAsnGluLysPheSerPhe 180
QY 922 TATAGTAAATTTAGTCTCTGGAGATGGGAGCTATAGATGCTTAAGAGCTTAAAGGTTCAA 981
```


Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAAGCTTTGTCTCTTCCAGAAATATCTGCTCAAGCTGATGGGAGCTTCT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGlnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACCAAGTTCTCTGCTATGCTTAACGAGGCTCTATTCCCTTTGTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTCAGGAGTAGAAGGGGAGGATTGCTGCTGTTTCAGATGGGAGGAGGAGTGCA	1161
Db	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAGTTTCCAGAAATACGCGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGCAACGTAGCCGAGTAGGAGGAGGATTACTCTACGGGAACGTGCTTTCCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAACCTTTGTTCTCAACAATGTTGCTTCTCTCTTTACATGCTGCTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu	320
QY	1342	CAACCAACAGCTGGACAGGCTTCTTAATACAGTAATAATTACGAGATGGAGAGTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle	340
QY	1402	TTCTGTAAGAATCGTCGCAAA---GCAGGATCAATAACTCTCGATCAGTTTCTTTGAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGGAGTAGTTTTCTTTAGTACATGTAGCTGCTGGAAAGGGGAGCTATTTAT	1518
Db	361	GlyGluGlyValValPhePheSerAsnValAlaAlaGlyLysGlyAlaIleTyr	380
QY	1519	GCACAAAGCTTCGGTTGCTACTGCGCCCTGTACAAATTTTAAGGATATCGCTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGGTGGACGATTATTTAGAGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGGA	1638
Db	401	AspGlyGlyAlaIleTyrLeuGlyLysSerGlyLysLeuSerLeuSerAlaAspTyrGly	420
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCGTAACGTGCTCCTCAAGCCATTTTCATGGGATCGGAGGAGGAAATAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
QY	1759	AGAGCTAAACAGCGGATCAGATCTCTTAAATGATCCCATCGAGATGGCAACGGAAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACCAGCCAGCGAGTCTTCCAAACTCTCAAAATTAACGATGCTGAAGGATACACAGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly	500
QY	1879	GATATTGTTTGTAAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAAGGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATTTCTTCTGTAAGGCAAAATATTCAGTGAATCTCTAAGTCACAGAGTGGG	1998
Db	521	ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATGGAAGCTGGAGTACATGGGATTTGTAACTCCCAACCAACCAACAG	2058
Db	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTCGCGCTAAATCAGTTGATCAGCTTTCCTCAATCTGCATTTCTCTTCTTTCTTG	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
QY	2119	TTAGCAAAACAATGCAAGTTACGAATCTCTTACCAATCTCTCCAGCCAGCAAGATCTCATCT	2178
Db	581	LeuAlaAsnAsnAlaValThrAsnProThrAsnProAlaGlnAspSerHisPro	600
QY	2179	CGAGTCATTGGTAGCAACACTGCTGTTCTGTACAAATAGTGGGCCCTATCTTTTTCAG	2238
Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTTCGATGATACAGCTTATGATAGGTATGATTCGCTTAATCAAAAATCAAT	2298
Db	621	AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTTCCTGAATTAACAGTTAGGAGTAAGCCCCAGCTTAATGCCCCATCAGATTTGACTTA	2358
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATCCCTAAGTATGCTATCAAGGAAGCTGGAAGCTGCGTGGGATCTTAAT	2418
Db	661	GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrPheLysLeuAlaIleThrAspProAsn	680
QY	2419	ACAGCAAAATAATGCTCTTATATCTCTGAAAGCTACATGAGCTAAAACTCGGTATAATCTCT	2478
Db	681	ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro	700
QY	2479	GGGCTGACGAGTAGTCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538
Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle	720
QY	2539	CGATCTGCGCATTCAGCAATTCAGCAAGTGGGATGGCGCTCTTATTGTCGAGGATTA	2598
Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740
QY	2599	TGGGTTCTGCGAGTTTCGAATTTCTTCTATCATGACCGCGATGCTTTTAGTCTCAGGATAT	2658
Db	741	TyrValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760
QY	2659	CGGTATATTAGTGGGGTTTATCTCTTAGAGCAAACTCTTCTTTCGATCATCGATGTTT	2718
Db	761	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780
QY	2719	GGTCTAGCATTTACCAAGTATTGCTAGATCTTAAAGATTATGCTAGTGTGCTGCTCAAT	2778
Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn	800
QY	2779	CATCATGCTTGCATAGGATCCGTTTATCTATCTTACCAACAAGCTTTTATGTGGATCCTAT	2838
Db	801	HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr	820
QY	2839	TTGTTCCGAGATCCGTTTATCCGTGCTAGTACGGGTTGGGATCAGCATATGAAAACC	2898
Db	821	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840
QY	2899	TCATATACATTTCCAGAGGAGCGATGTTGCTGGGATAAATCTGCTGGCTGGAGAG	2958
Db	841	SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu	860
QY	2959	ATTGGAGCGGATTACCGATTGTGATTCTTCTTACCTCAAGCTCTATTGTAATGAGTTGCGT	3018
Db	861	IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg	880
QY	3019	CCTTTCGTCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGGCGAT	3078
Db	881	ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp	900
QY	3079	CAAGCTCGGCATTCAGAGCGGACATCTCTTAATCTATCAGTTCTGCTGGTGGAGTAAG	3138
Db	901	GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys	920

QY 3139 TTGTGATGATGTTCTAGTACACATCTCTAATAAATATAGCTTATGCGCGCTTATATCTGT 3198
Db |||||||
QY 921 PheAspArgCysSerThrHisProAsnLysIleSerPheMetGlyAlaTyrIleCys 940
Db |||||||
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAGCGCTCTATCCCATCAAGAGACATGG 3258
Db |||||||
QY 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Db |||||||
QY 3259 ACAACAGATGCTCTTTCATTTAGCAAGACATGGAGTGTGGTTPAGAGGATCTATGATGCT 3318
Db |||||||
QY 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Db |||||||
QY 3319 TCTCTAACAGATTAATATAGAGTATATGCGCATGGAAGATATGATATGAGATCGGATCTTCT 3378
Db |||||||
QY 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Db |||||||
QY 3379 CGAGGCTATGTTTTCAGTGCAGAGTATAGATGAGTCCGGTTC 3417
Db |||||||
QY 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
Db |||||||

RESULT 12

ADW29027

ID ADW29027 standard; protein; 1013 AA.

XX AC ADW29027;

XX DT 07-APR-2005 (first entry)

XX DB C_trachomatis polymorphic membrane protein (PmpG) Seq 21.

XX KW immunogenicity; antigen; vaccine; chlamydia trachomatis infection;

XX KW antibacterial; gene therapy.

XX OS Chlamydia trachomatis.

XX PN WO200502619-A2.

XX PD 13-JAN-2005.

XX PF 25-JUN-2004; 2004WO-US020491.

XX PR 26-JUN-2003; 2003GB-00015020.

XX PR 25-AUG-2003; 2003US-0497649P.

XX PR 02-FEB-2004; 2004GB-00002236.

XX PR 01-JUN-2004; 2004US-0576375P.

XX PA (CHIR) CHIRON CORP.

XX PI Grandi G, Finco O, Ratti G, Bonci A;

XX DR WPI; 2005-075653/08.

XX PT An immunogenic composition for preventing or treating Chlamydia
XX PT trachomatis infections comprises a combination of C. trachomatis
XX PT antigens, such as PepA, LcrE, ArtJ, DnaK, CT398, OmpH-like, L7/L12, OmcA,
XX PT or AtoS.

XX PS Example 1; SEQ ID NO 21; 110pp; English.

XX CC This invention relates to novel immunogenic compositions that comprise a
XX CC combination of Chlamydia trachomatis (C. trachomatis) antigens.
XX CC Specifically, it refers to a first antigen group consisting of PepA,
XX CC LcrE, ArtJ, DnaK and CT398 and a second antigen group consisting of PepA,
XX CC LcrE, ArtJ, DnaK, CT398, OmpH-like, L7/L12, OmcA, AtoS, CT547, Enolase,
XX CC HtrA and MurG. The present invention describes the development of a
XX CC vaccine and antibodies that recognize the immunogenic composition in
XX CC order to provide a method of neutralizing a C. trachomatis infection in a
XX CC mammal through raising an immune response. Accordingly, the
XX CC pharmaceutical compositions derived thereof exhibit antibacterial
XX CC activity and the group of antigenic DNA sequences can also be used in
XX CC gene therapy. This polypeptide sequence is a Chlamydia trachomatis
XX CC antigen of the invention.

SQ Sequence 1013 AA;
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 9 Gaps: 1
US-10-701-844-1 (1-4435) x ADW29027 (1-1013)
QY 382 ATGCAAGCTCTTCCATAGCTTCTTCTTCAATGATCTAGCTTATCTTGTGCTCT 441
Db |||||||
QY 1 MetGlnThrSerPheHisLeuPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
Db |||||||
QY 442 TTAATGGGGGGGATATGTCAGCAGAAATCATGCTTCTCAAGGAATTTACGATGGGAG 501
Db |||||||
QY 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
Db |||||||
QY 502 ACGTTAACTGATCATTTCCCTATCTATCTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db |||||||
QY 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Db |||||||
QY 562 TCTGCAGGAGATTAACATTAATAAATCTTGACAACTTATTCGAGCTTTCGCTTAACT 621
Db |||||||
QY 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
Db |||||||
QY 622 TGTTTTGGAACTTATTAGGGAGTATTACTGTTTAGGGAGGAGGACACTCGTTGACTTTC 681
Db |||||||
QY 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Db |||||||
QY 682 GAGAACATACGAGCTTCTCAAAATGGGCGACTCTAAGTAATAGCGTCTGATGACTG 741
Db |||||||
QY 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Db |||||||
QY 742 TTTACTATTGAGGGTTTAAAGAAATATCTCTTTCCAATTGCAATTGATTTACTTCCGCTA 801
Db |||||||
QY 121 PheThrIleGlyGlyPheLysGlyLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Db |||||||
QY 802 CTGCCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGAGGACAACTACACCGTCT 861
Db |||||||
QY 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
Db |||||||
QY 862 AATGTAATAATTTATTTCTAAACAGACTCTTTGTTTACTCAATAATAGAGAGTTCTCATTC 921
Db |||||||
QY 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlyLysPheSerPhe 180
Db |||||||
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTAGATGCTTAAGAGCTTAACGGTTCAA 981
Db |||||||
QY 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
Db |||||||
QY 982 GGAATTAGCAAGCTTTGTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db |||||||
QY 201 GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyAlaCys 220
Db |||||||
QY 1042 CAAAGTAGTCACCAAGTTTCTCTGCTATGCTTAACGAGGCTCTTATTCCTTTTAGCAAT 1101
Db |||||||
QY 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Db |||||||
QY 1102 GTTGAGGAGTAGAGGGGGAGGATTGCTGCTTTCAGGATGGGCGAGGAGGTGCTCA 1161
Db |||||||
QY 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
Db |||||||
QY 1162 TCATCTTACTTCAACAGAAGATCCAGTAGTAAGTTTTCAGAAATACTCGCGGTAGAGTTT 1221
Db |||||||
QY 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Db |||||||
QY 1222 GATGGGAAGTAGCCCGAGTAGAGGAGGATTTACTCTACGGGAACGTTGCTTTCCCTG 1281
Db |||||||
QY 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Db |||||||
QY 1282 AATAATGGAAAAACCTTGTCTTCTCAACAATGTTGCTTCTCTCTTACATGCTGCTAAG 1341

Db 301 AenAenGlyLeuThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaGlu 320
QY 1342 CAACCAACAAAGTGGAGCAGCGCTCTTAATACGAGTAATAATACCGAGATGGAGGACTATC 1401
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAenAenThrGlyAenGlyAlaIle 340
QY 1402 TTCTGTAAAGTGGTCGGCAA--GCAGGATCCAAATACCTCGGATCAGTTTCCTTTGAT 1458
Db 341 PheCysLeuAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAen 360
QY 1459 GGAGGGGAGTAGTTCTTTTCTAGTACAAATGTAGCTGCTGGGAAAGCGGAGACTATTTAT 1518
Db 361 GlyGluGlyValPhePheSerAenValAlaAlaGlySerGlyAlaIleTyr 380
QY 1519 GCCAAAAGCTCTCGGTGCTTAACCTGGCCCTGTACAATTTTAAAGGAATATCGCTAAT 1578
Db 381 AlaLeuLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
QY 1579 GATGGTGGAGCGATTTATTAGGAGAACTCGGAGACTCAGTTTATCTGCTGATTTATGGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyLeuLeuSerLeuSerAlaAenTyrGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAAACAGCCAAAGAGAAATGCTGCCGATGTAAT 1698
Db 421 AspIleIlePheAenGlyAenLeuLysArgThrAlaLysGluAenAlaAenValAen 440
QY 1699 GCGTAACTGTGCTCACAAAGCCATTTTCGATGGGATCGGAGGGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGGCAATCAGATCTCTTAAATGATCCCATCGAGATGGCAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAenProIleGluMetAlaAenGlyAen 480
QY 1819 AACAGCCAGCGACTCTCCAAACTCTTAAATTAACCATGCTGAAGGATACACAGGG 1878
Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAenGlyGlyLeuThrGly 500
QY 1879 GATATTGTTTTGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGGA 1938
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyrGlnAenValThrIleGluGlnGly 520
QY 1939 AGGATGTTCTTCGTGAAAGGCAAAATTAATCAGTGAATCTCTAAGTCAAGAGGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTCTAACTCCACAAACACACACACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAenPheValThrProGlnProGlnGln 560
QY 2059 CCTCTCCCGCTAATCAGTTGATCACGCTTTCCAAATCTGCAATTTGTCTCTTTCTTTTG 2118
Db 561 ProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAACAATGACGATTACGAATCTCTCCATCCCAATCCTCAGCGCGAAGATCTCATCCT 2178
Db 581 LeuAlaAenAenAlaValThrAenProProThrAenProProAlaGlnAenSerHisPro 600
QY 2179 GCAGTCATTGTTAGCACAACTGCTGTTCTGTACAAATAGTGGGCTATCTTTTGTAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATTTGGCTAGGTCTTAATCAAAAAATCAAT 2298
Db 621 AspLeuAenAenProThrAlaTyrAspArgTyrAspIrpLeuGlySerAenGlnLysIleAen 640
QY 2299 GTCTGGAATTAACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAenLeuThrLeu 660
QY 2359 GGGATGAGATGCTTACGATGCTATCAAGGAGCTGGAAGCTTGGTGGGATCCTTAAT 2418

661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTyrIleLysLeuAlaTyrAspProAen 680
2419 ACAGCAAAATATGTCCTTATCTACTGAAGCTACATGACTATAAATCTGGGTATATCTCT 2478
Db 681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAenPro 700
QY 2479 GGGCTCGAGCGAGTAGTCTTTTGGTTCCTCAATAGTATTTATGGGATCCATTTTACATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTyrGlySerIleLeuAenPhe 720
QY 2539 CQATCTGGCATTCAGCAATTCACCAAGTGTGGATGGCGCTCTTATTTGTGCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAenGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTCTATCATGACCGGATGCTTTAGTTCAGGAGATAT 2658
Db 741 TrpValSerGlyValSerAenPhePheTyrHisAenArgAenAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTTATCTTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTGCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAenTyrValValCysArgSerAen 800
QY 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCCAAACAGCTTTATGTGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACGGTTTGGGAATCAGCATATGAAAAACC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
QY 2899 TCATATACATTTTCAGAGAGAGCGATGTTGTTGGGATAATACTGTCTGCTGGCTGAGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrPheAenAenCysLeuValGlyGlu 860
QY 2959 ATGGAGCGGATTCAGGATTCGATTCTCCATCTAAGCTTATTTGATGATGTTGCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
QY 3019 CCTTTCTGTCAGCTCAGATTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAenHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCAATTCAGAGCGGACATCTCCTAAATCTATCATCTCTGTTGAGTCAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGCTACTGAGCAACGCTCCTATCCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTyr 960
QY 3259 ACAACAGATCCCTTTCATTTAGCAACATGAGGATGTTGTTGTTAGAGATCTATGATGCT 3318
Db 961 ThrThrAenAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAAATAGAGTATATGGCCATGAGAGATATGATGATCCAGATGCTTCT 3378
Db 981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAenThrSer 1000
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 13
AEAL9078
ID AEAL9078 standard; protein; 1013 AA.

XX AC AEA19078;
 XX DT 28-JUL-2005 (first entry)
 XX DE Chlamydia trachomatis protein - SEQ ID 81.
 XX KW chlamydia trachomatis infection; antibacterial; vaccine.
 XX OS Chlamydia trachomatis.
 XX PN US2005106162-A1.
 XX PD 19-MAY-2005.
 XX PF 16-DEC-2004; 2004US-00498327.
 XX PR 12-DEC-2001; 2001GB-00029732.
 XX PR 06-AUG-2002; 2002EP-00182233.
 XX PR 14-AUG-2002; 2001EP-00218924.
 XX PR 12-DEC-2002; 2002WO-IB005761.
 XX (GRAN/) GRANDI G.
 XX (RATT/) RATTI G.
 XX FI Grandi G, Ratti G;
 XX DR WPI; 2005-354777/36.
 XX DR N-PSDB; AEA19079.
 XX PT New immunogenic composition comprising Chlamydia trachomatis protein or its encoding nucleic acid, useful for immunizing against, or treating or preventing chlamydial infection.
 XX PS Claim 13; SEQ ID NO 81; 90pp; English.
 XX CC The invention comprises an immunogenic composition for the treatment or prevention of a Chlamydia trachomatis infection. The immunogenic composition contains a Chlamydia trachomatis protein or its encoding nucleic acid. The immunogenic composition of the invention is useful for treating or preventing chlamydial infection. The present amino acid sequence represents a Chlamydia trachomatis protein of the invention. CC NOTE: The present sequence is not shown in the specification, but was obtained from the USPTO website -
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
 XX SQ Sequence 1013 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5131.50 Matches: 983
 Percent Similarity: 98.8% Conservative: 18
 Best Local Similarity: 97.0% Mismatches: 11
 Query Match: 65.1% Indels: 1
 DB: 9 Gaps: 1
 US-10-701-844-1 (1-4435) x AEA19078 (1-1013)
 QY 382 ATGCAACGCTCTTCCATAAGTTCTTTCTTCAATGATCTAGCTTATTCTGCTGCTCT 441
 Db 1 MetGlnThrSerPheHisIlysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
 QY 442 TTAATGGGGGGGATATCGACAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
 Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
 QY 502 ACGTTAACTGATCATTTCCCTATACCTGTATAGGAGATCCGATGGGACTACTGTTTTT 561
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 QY 562 TCTGCAGGAGAGTTAACTATAAAAAATCTTGACAAATCTTATGACGCTTTGCTTTAAGT 621
 Db 61 SerAlaGlyGluLeuThrLeuIysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80

QY 622 TGTCTTGGAACTATTAGGAGTTTACTGTTTGGGAGGACACACTCGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 QY 682 GAGAACATACGACTCTTACAAATGGGCGACTCTTAAGTAATAGCGCTGCTGATGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 QY 742 TTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTACTTGCCTGA 801
 Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
 QY 802 CTGCTGCTGCAACGACTAATAAGGTTAGCAGACTCCGACGACACACTTACACCGTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
 QY 862 AATGTACTATTATTCTTAAACAGACTCTTTTGTGTACTCAATAATGAGAAGTTCTCATTC 921
 Db 161 AsnGlyThrIleTyrSerIysThrAspLeuLeuLeuLeuAsnGluIysPheSerPhe 180
 QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
 QY 982 GGAATTAGCAAGCTTTGTCTCTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
 Db 201 GlyIleSerIysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
 QY 1042 CAAGTAGTCACAGTTTCTCTGCTATGCTTAACGAGGCTCTTATGCTTGTAGCGAAT 1101
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
 QY 1102 GTTCAGAGTAAGAGGGAGGAGTCTGCTGTTTCAGATGGGACAGAGGAGTGTC 1161
 Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
 QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACCTGCGGTAGAGTTT 1221
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 QY 1222 GATGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACTGCTTTCTCTG 1281
 Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
 QY 1282 AATAATGGAAAACTTGTCTCAACAAATGTTGCTTCTCTGTTTACATGCTGCTGAAG 1341
 Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
 QY 1342 CAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATACGGAGATGGGAGGACTATC 1401
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
 QY 1402 TTCTGTAAAGATGGTGGCGCAA---GCAGGATCCATAACTCTGGAATCAGTTTCTTTGAT 1458
 Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 QY 1459 GGAGAGGAGTAGTGTCTTTTAGTAGCAATGTAGTCTGTGGGAAAGGGGAGGACTATTTAT 1518
 Db 361 GlyGluGlyValAlaPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
 QY 1519 GCCAAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
 Db 381 AlaIysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 QY 1579 GATGCTGAGCGCAATTTATTAGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGA 1638
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
 QY 1639 GATATATTTTCGATGGGAATCTTAAAGAAACAGCAAAGAGAAATGCTGCCGATGTTAAT 1698
 Db 421 AspIleIlePheAspGlyAsnLeuIysArgThrAlaLysGluAsnAlaAlaAspValAsn 440

1699 GCGCTAACTGTCCTCAAGCCATTCGATGGGATCGGAGGAAATAACACATTA 1758
DB GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrLeu 460
1759 AGAGCTAAAGCAGGGCATCAGATTCCTCTTTAAATGATCCCATCGAGATGCGAAACGGAAT 1818
DB ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
1819 AACAGCCAGCCAGCTCTCCAAACTCTAAAAATTAACGATGGTGAAGGATACACAGG 1878
DB AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyThrGly 500
1879 GATATTGTTTTCCTAAATCGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGCA 1938
DB AspIleValPheAlaAsnGlyAsnSerThrLeuThrGlnAsnValThrIleGluGlnGly 520
1939 AGGATTGTTCTTCGTGAAGGCAAAATTTATCAGTGAATTCCTAAGTCAGACAGGTGG 1998
DB ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
1999 AGTCTGATATGAAGCTGGAGTACATGGGATTTTGTAACTCCCAACACACACAG 2058
DB SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
2059 CCTCTGCCGCTAACTCAGTTGATCAGCTTTCCTCAATCTGCATTTGCTCTCTCTTTG 2118
DB ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
2119 TTAGCAAAACAATGCAGTACGAATCCTCTTACCATCTCCAGCGCAAGATTCATCCT 2178
DB LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
2179 GAGTCATTGTTAGCAACAATCTGCTTCTGTTACAAATAGTGGGCTATCTTTTGTAG 2238
DB AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
2239 GATTTGGATGATACAGCTTATGATAGTATGATTGGCTAGTTCTAATCAAAAATCAAT 2298
DB AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
2299 GTCCTGAAATTAACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCATCAGATTTGACTTA 2358
DB ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
2359 GGAATGAGATGCTTAAGTATGCTATCAAGGAGCTGGAAGCTGGGTCGATCCTAAT 2418
DB GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
2419 ACAGCAATTAATGCTCTTATCTCTGAAGCTACATGGACTAAACTGGGTATATCCT 2478
DB ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
2479 GGGCTTGAGCGAGTACTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
DB GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
2539 CGATCTCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTGTCGAGGATTA 2598
DB ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
2599 TGGGTTCTCGAGTTTCGATTTCTTCTATCATGACCGCATGCTTTTAGTCAGGGATAT 2658
DB TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
2659 CGGTATATTAGTGGGGGTATCTCTTAGGAGCAAACTCTTACTTTGATCATCGATGTTT 2718
DB ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGATGATGTCGTTCAAT 2778
DB GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
2779 CATCATGCTTGCATAGGATCGTTTATCTATCTACCCAAACAGCTTTATGTGGATCTAT 2838

801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
2839 TTGTTCCGAGATGCGTTTATCCGTGCTAGCTACGGTTTGGGAATCAGCATATGAAAAACC 2898
DB LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
2899 TCATATACATTTTCAGAGGAGGAGCATGTTTCGTTGGGATAATAACTGTCGTGGTGAAG 2958
DB SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
2959 ATTGAGCGGATTAACCATTTGATTAATCTCATCTTAAGCTCTATTGATGAGTTGGT 3018
DB IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
3019 CTTTTCGTCGACAGCTGAGTTTCTTATGCCATCATGAATCTTTTACAGAGGAAGCGCAT 3078
DB ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138
DB GlnAlaAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
3139 TTTGATCGATGTTCTAGTACATCCTTAATAATATAGCTTTATGGCGCTTATATCTGT 3198
DB PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
3199 GATGTTATCGCACCATCTCTGTTACTGAGCAACGCTCCTATCCATCAAGAGACATGS 3258
DB AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
3259 ACAACAGATGCTCTTCATTTAGCAAGACATGAGTTGTGGTTAGAGGATCTATGTATGCT 3318
DB ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
3319 TCTTAACAAGTAATATAGATATATGATGATGATGATGATGATGATGATGATGATGAT 3378
DB SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
3379 CGAGCTATGGTTTTCAGTTCAGGAGTACAGTCCGGTTC 3417
DB ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 14
AEAL9002 standard; protein; 1013 AA.
XX AEAL9002;
XX 28-JUL-2005 (first entry)
XX Chlamydia trachomatis protein - SEQ ID 5.
XX chlamydia trachomatis infection; antibacterial; vaccine.
XX Chlamydia trachomatis.
XX US2005106162-A1.
XX 19-MAY-2005.
XX 16-DEC-2004; 2004US-00498327.
XX 12-DEC-2001; 2001GB-00029732.
XX 06-AUG-2002; 2002EP-00182233.
XX 14-AUG-2002; 2001EP-00218924.
XX 12-DEC-2002; 2002WO-1B005761.
XX (GRAN/) GRANDI G.
XX (RATT/) RATTI G.
XX Grandi G, Ratti G;
XX

DR MPI; 2005-354777/36.
 DR N-F5DB; AEA19003.
 XX New immunogenic composition comprising Chlamydia trachomatis protein or
 PT its encoding nucleic acid, useful for immunizing against, or treating or
 PT preventing chlamydial infection.
 XX Claim 13; SEQ ID NO 5; 90pp; English.
 XX The invention comprises an immunogenic composition for the treatment or
 CC prevention of a Chlamydia trachomatis infection. The immunogenic
 CC composition contains a Chlamydia trachomatis protein or its encoding
 CC nucleic acid. The immunogenic composition of the invention is useful for
 CC treating or preventing chlamydial infection. The present amino acid
 CC sequence represents a Chlamydia trachomatis protein of the invention.
 CC NOTE: The present sequence is not shown in the specification, but was
 CC obtained from the USPTO website -
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
 XX
 SQ Sequence 1013 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5131.50 Matches: 983
 Percent Similarity: 98.8% Conservatives: 18
 Best Local Similarity: 97.0% Mismatches: 11
 Query Match: 65.1% Indels: 1
 DB: 9 Gaps: 1
 US-10-701-844-1 (1-4435) x AEA19002 (1-1013)
 QY 382 ATGCAACGCTCTTCATTAAGTCTTCTTCAATGATCTAGCTATCTTGCTGCTCT 441
 Db 1 MetGlnThrSerPheHisLeuSerMetIleLeuAlaIleValSerCysSer 20
 QY 442 TTAATATGGGGGATATGCACAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501
 Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
 QY 502 AGCTTAACCTGTATCTATCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTT 561
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 QY 562 TCTGACGAGAGTTAACTATAAAATCTTGACAAATCTATTTGCGAGCTTGTGCTTAAGT 621
 Db 61 SerAlaGlyGluLeuThrLeuLeuSerLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
 QY 622 TGTTTTGGGAATTTATAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 QY 682 GAGAACATACGACTCTTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 QY 742 TTTATCTATTGAGGTTTTAAAGAAATATCTCTTTTCCAAATGCAATTCATTACTTGGCGTA 801
 Db 121 PheThrIleGluGlyPheLeuGlyGluLeuSerPheSerAsnSerLeuLeuAlaVal 140
 QY 802 CTGCGCTGCTGCAAGCACTAATAAGGCTAGCCAGACTCCGACGACACATCTACACGCTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
 QY 862 AATGGTACTATTATTCTAAAACAGATCTTTGTTACTCAATAATGAGAAGTTCTCATTC 921
 Db 161 AsnGlyThrIleTyrSerIleThrAspLeuLeuLeuLeuAsnAsnGlyLeuSerPhe 180
 QY 922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTTAAGGCTTAACGGTTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLeuSerLeuThrValGln 200
 QY 982 GGAATTTAGCAAGCTTTGCTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTCT 1041

Db 201 GlyIleSerLeuValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
 QY 1042 CAAGTAGTCACAGTCTTCTGCTATGGCTAACGAGGCTCTATTTGCTTTGTAGCGAAT 1101
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
 QY 1102 GTTCGAGAGTAACAGAGGGGAGGATGCTGCTGTTCCAGGATGGCGACGAGGTGTCA 1161
 Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
 QY 1162 TCATCTACTCTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATACTGCGGTAGAGTTT 1221
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 QY 1222 GATGGAACTAGTCCCGAGTAGGAGGAGGATTTACTCTCTACGGGAAGTGTCTTCTG 1281
 Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
 QY 1282 AATAATGGAAAACTGTTTCTCAACAATGTTGCTCTCTCTCTGTTTACATTGCTGCTAAG 1341
 Db 301 AsnAsnGlyIleThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
 QY 1342 CAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
 QY 1402 TTCTGTAAAGATGTGCGCAA---GCAGGATCCAAATACTCTGATCAGTTCTCTTCTGAT 1458
 Db 341 PheCysLeuAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 QY 1459 GGAGAGGAGTAGTCTTCTTAGTAGCAATGCTAGTCTGGGAAAGGGGAGCTATTTAT 1518
 Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLeuGlyGlyAlaIleTyr 380
 QY 1519 GCCAAAAAGCTCTCGTTGCTCAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578
 Db 381 AlaLeuLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 QY 1579 GATGTGAGGAGTATTTATAGGAGATCTCGAGAGCTCAGTTTATCTGCTGATATATGA 1638
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGlyLeuSerGlyLeuSerLeuSerAlaAspTyrGly 420
 QY 1639 GATATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCGCGATGTTAAT 1698
 Db 421 AspIleIlePheAspGlyAsnLeuLeuArgThrAlaLeuGluAsnAlaAlaAspValAsn 440
 QY 1699 GGCGTAACCTGTCTCTCAAGCCATTTTCGATGGGATCGGAGGGGAAATAACGACATTA 1758
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
 QY 1759 AGAGCTAAAGCGGGCATCAGATCTCTTTAATGATCCCATCGAGATGCGAACCGAAT 1818
 Db 461 ArgAlaLeuAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
 QY 1819 AACCGACGAGCGAGCTCTTCCAACTCTTAAATAATTAACGATGGTGAAGGATACACAGGG 1878
 Db 481 AsnGlnProAlaGlnSerSerGluProLeuLeuIleAsnAspGlyGlyGlyTyrThrGly 500
 QY 1879 GATATTGTTTGTCTTAATGAAGCACTACTTTGTACCAAAATGTTTACGATAGACAGGA 1938
 Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
 QY 1939 AGGATTGTTCTCGTGAAGCAAAATATCAGTGAATTTCTAAGTCAGACAGGTGGG 1998
 Db 521 ArgIleValLeuArgGluLeuAlaLeuSerValAsnSerLeuSerLeuSerGlnThrGly 540
 QY 1999 AGTCGTATATGGAAGCTGGAGTACATGGGATTTGTAACTTCCCAACACCAACACAG 2058
 Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
 QY 2059 CCTCTGCGGCTTAATCAGTTGATCAGCTTTTCCAAATCTGCAATTTGCTCTTCTTTG 2118
 Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580

QY 2119 TTGACAAATGCGAGTACCAATCTCTACCAATCTCTCCAGCGCAAGATTCTCATCT 2178
 DB 581 LeuAlaAsnAlaValThrAsnProPheThrAsnProPheAlaGlnAspSerHisPro 600
 QY 2179 GCAATGATGAGTACCACTGCTGCTTCTGTACAAATAGTGGGCTATCTTTTGTAG 2238
 DB 601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePheGlu 620
 QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAATCAAT 2298
 DB 621 AspLeuAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnIleAsp 640
 QY 2299 GTCCTGAATACAGTTAGGACTAGCCCGAGCTAATGCCCATACAGATTGACTTA 2358
 DB 641 ValLeuIleGlySerGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
 QY 2359 GGAATGAGATGCTTAAGTATGGCTATCAAGGAGCTGGAAGCTTCGGTGGATCTAAT 2418
 DB 661 GlyAsnGluMetProIleGlyThrGlyGlnGlySerTyrIleLeuAlaTyrAspProAsn 680
 QY 2419 ACAGCAAAATAGTCTTATCTGCTGAAAGCTACATGAGCTAAACTGGGTATATCTCT 2478
 DB 681 ThrAlaAsnAsnGlyProTyrThrLeuIleAlaThrTyrThrIleSerGlyTyrAsnPro 700
 QY 2479 GGGCTGAGGAGTACTCTTCTGCTTCCAAATAGTTATGGGATCCATTTAGATATA 2538
 DB 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrIleLeuAspIle 720
 QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGGCTCTTATTGTCAGGATTA 2598
 DB 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
 QY 2599 TGGGTTCTGGAGTTTCCGATTTCTTCTATCATGACCGGATGCTTAGTCAAGGATAT 2658
 DB 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
 QY 2659 CGGTATATTAGTGGGGTTATCTCTAGGAGCAACTCTTCTGATCATCGATGTTT 2718
 DB 761 ArgTyrIleSerGlyGlySerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
 QY 2719 GGTCTAGCATTTACCAAGTATTTGGTAGATCTAAAGATTATGTAGTGTCTGCTCAAT 2778
 DB 781 GlyLeuAlaPheThrGluValPheGlyArgSerIleAspTyrValValCysArgSerAsn 800
 QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTACCAACAGCTTTATGTGATCTAT 2838
 DB 801 HisHisAlaCysIleGlySerValTyrLeuSerThrIleGlnAlaLeuGlySerTyr 820
 QY 2839 TTGTTGGAGATGCGTTTATCCGTGTAGCTACGGTTTGGGATCGGATATGAAAC 2898
 DB 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
 QY 2899 TCATATACATTCAGAGGAGCGATGTTCTGGGATTAATCTCTGCTGGGAGAG 2958
 DB 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
 QY 2959 ATTGGAGCGGATTTACCGATTGTGATTACTTCCATCTAAGCTCTATTGTAATGATGCGT 3018
 DB 861 IleGlyValGlyLeuPheProIleValIleThrProSerIleLeuTyrLeuAsnGluLeuArg 880
 QY 3019 CTTTTCGTGCAAGCTGATGTTTCTTATGCGCATCATCAATCTTTTACAGAGGAGCGAT 3078
 DB 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
 QY 3079 CAAGCTCGGATTCAGAGCGGACATCTCTCAATCTATCAGTTCTCTGTTGGAGTCAAG 3138
 DB 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValIys 920
 QY 3139 TTTGATCGATGTTCTAGTACATCTCTAAATAATATAGTCTTATGGCGGCTTATCTCTGT 3198
 DB 921 PheAspArgCysSerSerThrHisProAsnIleIleSerPheMetGlyAlaTyrIleCys 940

QY 3199 GATGCTTATCGACCATCTCTGGTACTGAGACAAGCGTCTCTATCCATCAAGAGCATGG 3258
 DB 941 AspAlaTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTyr 960
 QY 3259 ACAACAGATGCTTTTCATTATAGCAAGACATCGAGTTGTGGTTAGAGGATCTATGATGCT 3318
 DB 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 QY 3319 TCTCTAAACAAGTAAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATCTTCT 3378
 DB 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 QY 3379 CGAGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417
 DB 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
 RESULT 15
 AEA19080
 ID AEA19080 standard; protein; 1013 AA.
 XX AEA19080;
 XX 28-JUL-2005 (first entry)
 XX Chlamydia trachomatis protein - SEQ ID 83.
 XX Chlamydia trachomatis infection; antibacterial; vaccine.
 XX Chlamydia trachomatis.
 XX US2005106162-A1.
 XX 19-MAY-2005.
 XX 16-DEC-2004; 2004US-00498327.
 XX 12-DEC-2001; 2001GB-00029732.
 XX 06-AUG-2002; 2002EP-00182233.
 XX 14-AUG-2002; 2001EP-00218924.
 XX 12-DEC-2002; 2002WO-1B005761.
 XX (GRAN/) GRANDI G.
 XX (RATT/) RATTI G.
 XX Grandi G, Ratti G;
 XX WPI; 2005-354777/36.
 XX N-PSDB; AEA19081.
 XX New immunogenic composition comprising Chlamydia trachomatis protein or
 its encoding nucleic acid, useful for immunizing against, or treating or
 preventing chlamydial infection.
 XX Claim 13; SEQ ID NO 83; 90pp; English.
 XX The invention comprises an immunogenic composition for the treatment or
 prevention of a Chlamydia trachomatis infection. The immunogenic
 composition contains a Chlamydia trachomatis protein or its encoding
 nucleic acid. The immunogenic composition of the invention is useful for
 treating or preventing chlamydial infection. The present amino acid
 sequence represents a Chlamydia trachomatis protein of the invention.
 XX NOTE: The present sequence is not shown in the specification, but was
 obtained from the USPTO website -
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
 XX
 SQ Sequence 1013 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5125.50 Matches: 980
 Percent Similarity: 98.8% Conservative: 21
 Best Local Similarity: 96.7% Mismatches: 11
 Query Match: 65.0% Indels: 1

DB:	9	Gaps:	1
US-10-701-844-1 (1-4435) x ABA19080 (1-1013)			
QY	382	ATGCAGAGCTCTTCCATAAGTCTCTTCTCAATGATTCAGCTTATTCTTGCTGCTCT	441
Db	1	MetGlnThrSerPheHisLeuPheLeuSerMetLeuLeuAlaIleTySerCysSer	20
QY	442	TTAAATGGGGGGGATATGCAGCAGAAATCATGTTCTCTCAAGGAATTTACGATGGGGAG	501
Db	21	LeuSerGlyGlyTyAlaAlaGluIleMetIleProGlnGlyIleTyAspGlyGlu	40
QY	502	ACGTTAACTGATCATTTCCCTATCTCTTATAGGAGATCCGAGTGGGACTACTGTTTTT	561
Db	41	ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGCAGCAGAGTTAAACATTAATAAATCTTCACAAATCTATTGACGCTTTCCTTTAAGT	621
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
QY	622	TGTTTTGGAACTTATTAGGGAGTTTTACTGTGTTTGGGAGAGGACACTCGTTGACTTTC	681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyA-gGlyHisSerLeuThrPhe	100
QY	682	GAGAACATACCGACTTCTACAAATGGGCGAGCTCTTAAGTAATAGCGCTGCTGATGGACTG	741
Db	101	GluAsnIleA-gThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTAGGGTTTTAAGAATATCTCTTCTCAATTCGAATTCATTTACTTCCCGTA	801
Db	121	PheThrIleGluGlyPheGlyGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGCTGCACCACTTAATAAGGTAGCCAGACTCCGACGACACACATCTACACCGTCT	861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGTTACTATTATTCTTAAACAGATCTTTTGTCTCAATTAATAGAGAGTTCTCATTC	921
Db	161	AsnGlyThrIleTySerLysThrAspLeuLeuLeuLeuLeuLeuLeuLeuLysPheSerPhe	180
QY	922	TATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGACTTAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAAATAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGlnLeuThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACCAAGTTCTCTGCTATGCTTAACGAGGCTCTTATGCTTGTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTCGAGGATGAAGGGGGAGTGTCTGCTGTTCAGGATGGCGCAGCGAGGTGCA	1161
Db	241	ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGCGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGGAACGTACCGCGAGTAGGAGGAGTTTACTCTCAGGGAAGCTCTCTTTCCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAAACTTTCTCAACAATGTTGCTTCTCCCTGTTTACATTCGCTGTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu	320
QY	1342	CAACCAACAGGTGAGCAGGCTTCTAATACAGGTAAATAATACGAGATGAGGAGCTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle	340
QY	1402	TTCTGTGAAGAATGGTGGCAA---GCAGATCCAAATAACTCTGGATCAGTTCTCTTTGAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGGAGGAGTAGTTTTCTTTTAGTACCAATGTAGCTCTGGGAAAGGGGAGCTATTAT	1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTy	380
QY	1519	GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATCGTCGAGCGATTTATTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA	1638
Db	401	AspGlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyGly	420
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAGAAATGTCGCGATGTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCTAACTGTCTCTCACAGCCATTTTCGATGGGATCGGAGGAGGAAATTAACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
QY	1759	AGAGCTAAAGCAGGCGCATCATCTCTTAAATGATCCATCGAGATCGGCAACCGAAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACGACGACGCGAGTCTTCCAAACTCTTAAAAATTAACGATGGTGAAGGATACACAGGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly	500
QY	1879	GATATTGTTTGTCTAATGGAAGCAGTACTTTGTATCCAAATGTTTACGATAGACCAAGGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATGCTCTCTCTGAAAGGCAAAATTAATCAGTGAATCTCTAAGTCACAGAGTGGG	1998
Db	521	ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACAG	2058
Db	541	SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCAATTTGCTCTTCTTCTCTTG	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
QY	2119	TTAGCAACAATGCAGTTACCAATCTCTTACCAATCTCCAGGCGCAAGATTTCTCATCT	2178
Db	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	GCAGTCATTGTTAGCACCAACTGCTGTTCTGTACAATTAGTGGGCTTATCTTTTTGAG	2238
Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTGGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2298
Db	621	AspLeuAspAspThrAlaTyAspArgTyAspTyLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTCTCGAAATACAGTTAGGAGTAAAGCCCGGCTAATGCGCCCATCAGATTTGACTTA	2358
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2418
Db	661	GlyAsnGluMetProLysTyGlyTyGlnGlySerTyLeuLysLeuAlaTyAspProAsn	680
QY	2419	ACAGCAAAATATGTCCTTATCTCTGAAAGCTACATGCACTAAACATGGGTATATCTCT	2478
Db	681	ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTyThrLysThrGlyTyAsnPro	700
QY	2479	GGGCTGAGCAGTACGCTTCTTGGTTCCAATAGTTTATGGGATCCATTTTAGATATA	2538

Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCATTCAGCAATTCACGAAGTGTGGATGGCGCTCTTATTGTGAGGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTCTATCATGACCGCATGCTTTAGTTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTATTCTCTTAGGAGCAAACTCTCTCTTGGATCATGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGATAGGATCCGTTTATCTATCTACCAACAGCTTTATGTGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTTCGGAGATCGCTTTATCCGCTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTCCGAGAGAGCGATGTTCTGTTGGATAATACTGTCTGGCTGGAGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTCAGGATTCGATCTCTCATCTTAAGCTCTTATTTGAATGATGTCCT 3018
Db 861 IleGlyValGlyLeuProIleThrValIleThrProSerLysLeuTyrLeuAsnGlnLeuArg 880
QY 3019 CTTTTCGTCGAAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCATTCAGAGCGGACATCTCTTAATCTATCATGTTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTGTATCGATGTTCTAGTACATCTCTAATATATATAGTCTTTATGGCGCTTATCTCT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGTGCTAGTACAGACAACTCTCTATCCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATCCCTTCATTAGACAGATGGAGTTGGTTAGAGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAATATAGNAGTATGCGCATGGAGATGAGTATCAGATCGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgGlyTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 16
ID AEA19006
AC AEA19006 standard; protein; 1013 AA.
XX AEA19006;
AC AEA19006;
XX
DT 28-JUL-2005 (first entry)
XX
DE Chlamydia trachomatis protein - SEQ ID 9.
XX
XX chlamydia trachomatis infection; antibacterial; vaccine.

XX Chlamydia trachomatis.
OS US2005106162-A1.
PN 19-MAY-2005.
PD 16-DEC-2004; 2004US-00498327.
PF 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002EP-00182233.
PR 14-AUG-2002; 2001EP-00218924.
PR 12-DEC-2002; 2002WO-IB005761.
XX (GRAN/) GRANDI G.
PA (RATT/) RATTI G.
PI Grandi G, Ratti G;
XX WPI; 2005-354777/36.
DR N-PSDB; AEA19007.
XX New immunogenic composition comprising Chlamydia trachomatis protein or its encoding nucleic acid, useful for immunizing against, or treating or preventing chlamydial infection.
PS Claim 13; SEQ ID NO 9; 90pp; English.
XX The invention comprises an immunogenic composition for the treatment or prevention of a Chlamydia trachomatis infection. The immunogenic composition contains a Chlamydia trachomatis protein or its encoding nucleic acid. The immunogenic composition of the invention is useful for treating or preventing chlamydial infection. The present amino acid sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX SQ Sequence 1013 AA;
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5123.50 Matches: 979
Percent Similarity: 98.8% Conservative: 22
Best Local Similarity: 96.6% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 9 Gaps: 1
US-10-701-844-1 (1-4435) x AEA19006 (1-1013)
QY 382 ATGCAACGCTCTTCCATTAAGTCTTCTTCAATGATTCTAGCTTATTCTTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAAATGGGGGGGATATGACAGCAAAATCATGGTCTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACCTTAACCTGTATCATTTCCCTATCTGTATTAGGAGATCCGAGTGGGACTACTGTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGACGAGAGTTAACTTAAATCTTGCAATTCATTCTTATTCACGCTTTCCTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaAlaLeuProLeuSer 80
QY 622 TGTGTTGGGAACCTTATTAGGAGTCTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGCACTTCTACAAATGGGCGAGCTCTTAAGTAATAGCCGCTGCTGAGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120

QY	742	TTTACTATTAGGGTTTTAAAGAATTATCTCTTTTCCAAATTCAAATTCATTAATCTGCGCTA	801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGTGTCACAGACTAATAAGGGTAGCCAGACTCCGACGACACACATCTACACCGTCT	861
Db	141	LeuProAlaLeaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer	160
QY	862	AATGGTACTATTATTCTTAAACAGACTCTTTGTTACTCAATAATGAGAAGTCTTCATTC	921
Db	161	AsnGlyThrIleTySerIysThrAspLeuLeuLeuLeuAsnAsnGluTyPheSerPhe	180
QY	922	TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA	981
Db	181	TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT	1041
Db	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCACCAAGTTCTCTGCTATGCTAACGAGGCTCTATTGCTTTGTAGCGAAT	1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTGAGGAGTAAGAGGGGAGGATTGCTGCTGTTTCAGGATGGGCAGCAGGAGTGTC	1161
Db	241	ValAlaGlyValArgGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTTCACAGAGATCCAGTAGTAGTTTTCAGNAATCTACGCGTAGAGTTT	1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGGAAGTAGCCGAGTAGGAGGAGGATTACTCTACGGGAAAGTTCCTTCCTG	1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTySerTySerTySerTySerTySerTySer	300
QY	1282	AATAATGGAACCACTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTCGCTGAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyLeuAlaAlaGlu	320
QY	1342	CAACCAACAGATGGACAGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC	1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTySerGlyAspGlyGlyAlaIle	340
QY	1402	TTCTCTAAGAAATGGGCGCAA---GCAGGATCCAATACTCTGGATCAGTTTCCTTCAT	1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT	1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTy	380
QY	1519	GCCAAAAGCTCTCGGTGCTTAACCTGGGCCCTGTACAAATTTTAAAGAAATTCCTAAT	1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGCTGGAGCGATTTATTAGGAGACTCGAGAGCTCAGTTTATCTGCTGATTATGGA	1638
Db	401	AspGlyGlyAlaIleTySerLeuGlyGlySerGlyGlyLeuLeuSerLeuSerAlaAspTy	420
QY	1639	GATATTATTTTCGATGGGAATCTTAAAGAAACAGCAAGAGAATGCTGCCGATCTTAAT	1698
Db	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGCGTAACCTGCTCTCAACAGCCATTTTCGATGGGATCGGAGGGAATAATACGACATTA	1758
Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu	460
QY	1759	ACAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGATGCGCAACCGAAT	1818
Db	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACCAGCCAGCGAGTCTTCCAAACTTCTAAAAATTAACGATGCTGAAGGATACACAGG	1878
Db	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly	500
QY	1879	GATATTGTTTGTGTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGGA	1938
Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGlnGlnGly	520
QY	1939	AGGATTGTTCTTCGTGAAAAGGCAAAATATTCAGTGAATTCCTAAAGTCACAGGTGG	1998
Db	521	ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAACCAACCAACAG	2058
Db	541	SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CTCTCTCGCGTAATCAGTTGATCACGCTTTCCAATCTGCAATTTGCTCTTCTCTCTTG	2118
Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu	580
QY	2119	TTAGCAACCAATGAGTTCACGAATCTCTCAACCAATCTCCAGCGCAAGATTCTCATCT	2178
Db	581	LeuAlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	GCAGTCAATGTTGACACAACTGCTGTTTCAATTAATTTAGTTGGGCTATCTTTTTCAG	2238
Db	601	AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTGATGATGATCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT	2298
Db	621	AspLeuAspAspThrAlaTyAspArgTyAspIleLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTCCTGAAAATPACAGTTAGGACTAAAGCCCCAGCTAATGCCCCATCAGATTTGACTCTA	2358
Db	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGTGCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCCTAAT	2418
Db	661	GlyAsnGluMetProLysTyGlyTyGlnGlySerTyLysLeuAlaTrpAspProAsn	680
QY	2419	ACAGCAATATGCTCTTATCTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT	2478
Db	681	ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTyThrLysThrGlyTyAsnPro	700
QY	2479	GGGCTGAGCAGTAGCTTCTTGTGTTCCAAATAGTTTATGGGATCCATTTTAGATATA	2538
Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyLysIleLeuAspIle	720
QY	2539	CGATCTGGCAATTCAGCAATTCAGCAAGTGTGGATGGGCGCTCTTATGTGCGAGGATTA	2598
Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu	740
QY	2599	TGGGTTTCTGAGTTTCGAATTTCTCTCATCATGACCGCGATGCTTTTAGGTCAGGATAT	2658
Db	741	TrpValSerGlyValSerAsnPheTyThrHisAspArgAspAlaLeuGlyGlnGlyTy	760
QY	2659	CGGTATATTTAGTGGGGTTTATCTCTTAGGAGCAAACTCCTACTTTTGGATCATCGATGTT	2718
Db	761	ArgTyIleSerGlyTySerLeuGlyAlaAsnSerTyPheGlySerSerMetPhe	780
QY	2719	GGTCTAGCAATTCAGAAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAT	2778
Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsn	800
QY	2779	CATCATGCTCCATAGGATCCGTTTATCTACTACCAACCAAGCTTTTATGTGATCCTAT	2838
Db	801	HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy	820
QY	2839	TTGTTTCGAGATGCGTTTATCCGCTAGCTACGCGTTTGGGAATCAGCATATGAAAACC	2898
Db	821	LeuPheGlyAspAlaPheLeuArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThr	840
QY	2899	TCATATACATTTGCGAGGAGCGATGTTGTTGGGATTAATACTGCTGCTGCGAGNG	2958

||||| SerThrPheAlaGluSerAspValArgTrpAspAsnCysLeuValGlyGlu 860
Db
Qy 2959 ATTGAGCGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGAATGAGTTGGT 3018
XX
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuThrLeuAsnGluLeuArg 880
Qy 3019 CTTTCGTGCAAGCTGAGTTTCTTATCCGATCATGAATCTTTTACAGAGGAGCGCAT 3078
Db
Db 881 ProPheValGlnAlaGluPheSerThrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGGCATTCACAGCGGACATCTCTAAATCTATCATGTTCTCTGGAGTGAAG 3138
Db
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCATGTTCTAGTACATCATCTTAATAATATAGCTTTATGGCGCTTATCTGT 3198
Db
Db 921 PheAspArgCysSerThrHisProAsnLysThrSerPheMetGlyAlaThrIleCys 940
Qy 3199 GATGCTTATCGACCATCTCTGTTACTGAGCAACGCTCTATCCCATCAAGAGCATGG 3258
Db
Db 941 AspAlaThrArgThrIleSerGlyThrGlnThrThrLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATGCTTTTCATTAGCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCT 3318
Db
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetThrAla 980
Qy 3319 TCTTAAACAAGTAATATAGATATATGCGCATGGAAGATATGATATCGAGATGCTTCT 3378
Db
Db 981 SerLeuThrSerAsnIleGluValThrGlyHisGlyArgThrGluThrArgAspThrSer 1000
Qy 3379 CGAGCTATGTTTTCAGTGCAGGAGTATGAGTCCGGTTC 3417
Db
Db 1001 ArgGlyThrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 17

AEA19090
ID AEA19090 standard; protein; 1013 AA.

AC AEA19090;

DT 28-JUL-2005 (first entry)

DE Chlamydia trachomatis protein - SEQ ID 93.

XX chlamydia trachomatis infection; antibacterial; vaccine.

KW Chlamydia trachomatis.

OS Chlamydia trachomatis.

XX US2005106162-A1.

PN 19-MAY-2005.

XX 16-DEC-2004; 2004US-00498327.

XX 12-DEC-2001; 2001GB-00029732.

PR 06-AUG-2002; 2002EP-00182233.

PR 14-AUG-2002; 2001EP-00218924.

PR 12-DEC-2002; 2002WO-IB005761.

XX (GRAN/) GRANDI G.

PA (RATT/) RATTI G.

XX Grandi G, Ratti G;

XX WPI: 2005-354777/36.

DR N-PSDB; AEA19091.

XX New immunogenic composition comprising Chlamydia trachomatis protein or

PT its encoding nucleic acid, useful for immunizing against, or treating or

PT preventing chlamydial infection.

XX Claim 13; SEQ ID NO 93; 90pp; English.

XX The invention comprises an immunogenic composition for the treatment or
CC prevention of a Chlamydia trachomatis infection. The immunogenic
CC composition contains a Chlamydia trachomatis protein or its encoding
CC nucleic acid. The immunogenic composition of the invention is useful for
CC treating or preventing chlamydial infection. The present amino acid
CC sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was
CC obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.

XX SQ Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5123.50 Matches: 979
Percent Similarity: 98.8% Conservative: 22
Best Local Similarity: 96.6% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 9 Gaps: 1

US-10-701-844-1 (1-4435) x AEA19090 (1-1013)

Qy 382 ATGCAACGCTCTTCCATAGTTCCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaThrSerCysCysSer 20
Qy 442 TTAATGCGGGGATATGACGAGAAATCATGTTCTCAAGGAATTCAGATGGGAG 501
Db 21 LeuSerGlyGlyGlyThrAlaAlaGluIleMetIleProGlnGlyIleThrAspGlyGlu 40
Qy 502 ACCTTAACGCTATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTCAGGAGAGTTAACTTAAATACTTGACAAATCTTGACAAATCTTGACAGCTTTGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80
Qy 622 TGTTTTGGAACTTATTAGGGAGTTTACTCTTTTAGGGAGGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACGACCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTACTATTGAGGTTTAAAGAAATATCTTTTCCAAATTTGCAATTCATTTCTGCGGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTGCTGCTGCACGACTAATAAGGGTAGCCAGACTCCGACGACACATCTACACCGTCT 861
Db 141 LeuProAlaAlaThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
Qy 862 AATGCTACTATTATTCTAAAACAGATCTTTTGTACTCAATAATGAGAAAGTTCTCATTC 921
Db 161 AsnGlyThrIleThrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
Qy 982 GGAATTAGCAGCTTTGTGCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTGCT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTCACAGTTCTCTGCTATGCTTAACGAGGCTCTTATGCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy 1102 GTTCAGGAGTAGAAGGGGAGGAGTTCGCTGTTTCTGAGATGGGAGGAGTGTCA 1161

Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAaspGlyGlnGlnGlyValSer 260
 QY 1162 TCATCTACTTCAACAGAAAGATCCAGTAGTAAGTTTTCAGAAATACTCGGGTAGAGTTT 1221
 Db 261 SerSerThrThrGluAaspProValValSerPheSerArgAanThrAlaValGluPhe 280
 QY 1222 GATGGGAACCTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTTCTCG 1281
 Db 281 AspGlyAasnValAlaAlaGValGlyGlyLeuLeuSerThrGlyAasnValAlaPheLeu 300
 QY 1282 AATAATCGAAAAACCTTGTTTCTCAACAATGTGCTTCTCTCTGTTTACATGCTGCTAAG 1341
 Db 301 AasnAangGlyLysThrLeuPheLeuAasnValAlaSerProValTyrLeuAlaAlaGlu 320
 QY 1342 CAACCAACAGTGCAGAGCTTCTAATACGAGTAATAATACCGAGATGCGAGGAGCTATC 1401
 Db 321 GlnProThrAangGlyGlnAlaSerAasnThrSerAaspAanTyrGlyAaspGlyAlaIle 340
 QY 1402 TTCTGTAAAGATGTCGCA--GCAGGATCCCAATAACTCTGGATCAGTTTCTCTTGAT 1458
 Db 341 PheCysLysAangGlyAlaGlnAlaAlaGlySerAasnAasnSerGlySerValSerPheAap 360
 QY 1459 GGAGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTCTCGGAAAGGGGAGCTATTAT 1518
 Db 361 GlyGluGlyValValPhePheSerSerAasnValAlaAlaGlyLysGlyAlaIleTyr 380
 QY 1519 GCCAAAAAGCTCTCGGTGCTAACTGTGGCCCTGTACATTTTAAAGAAATATCGCTAAT 1578
 Db 381 AlaLysLysLeuSerValAlaAasnCysGlyProValGlnPheLeuGlyAasnIleAlaAasn 400
 QY 1579 GATGGTGGAGGATTTATTTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
 Db 401 AspGlyGlyAlaIleTyrLeuGlyLysSerGlyGluLeuSerLeuSerAlaAaspTyrGly 420
 QY 1639 GATATTATTTTCGATGGGAATCTTAAAAAGAACAGCCAAAGAAATGTCGCGATGTTAAT 1698
 Db 421 AspilellePheAaspGlyAasnLeuLysArgThrAlaLysGluAasnAlaAaspValAasn 440
 QY 1699 GCGTAACGTGCTCTCAAGCCATTTCGATGGATCGGAGGGGAAATAACGACATTA 1758
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
 QY 1759 AGAGCTAAGCAGGCGATCAGATTCTCTTTAATGATCCATCGAGATGCGCAACCGAAAT 1818
 Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAasnAaspProIleGluMetAlaAasnGlyAasn 480
 QY 1819 AACCAAGCCAGCGAGCTCTTCCAACTCTTAAAAATTAAACGATGGTGAAGGATACACAGG 1878
 Db 481 AasnGlnProAlaGlnSerSerGluProLeuLysIleAasnAaspGlyGluGlyTyrThrGly 500
 QY 1879 GATATTGTTTTCGTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGCAAGA 1938
 Db 501 AspilelleValPheAlaAangGlyAasnSerThrLeuTyrGlnAasnValThrIleGlnGly 520
 QY 1939 AGGATGTTCTTCGTGAAGGCAAAATTTATCAGTGAATTCCTAAGTCAGACAGGTGG 1998
 Db 521 ArgIleValLeuArgGlnLysAlaLysLeuSerValAasnSerLeuSerGlnThrGlyGly 540
 QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACCAACCAACAG 2058
 Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAaspPheValThrProGlnProGlnGln 560
 QY 2059 CTCCTGCGCTAATCAGTTGATCAGCTTTCGATCTCGATTTGCTCTTCTCTCTTCTTCTTG 2118
 Db 561 ProProAlaAlaAangGlnLeuIleThrLeuSerAasnLeuHisLeuSerLeuSerLeu 580
 QY 2119 TTAGCAAAACATGACGTACGATCTCTCTACCAATCTCCAGCGCAAGATCTCATCCT 2178
 Db 581 LeuAlaAasnAlaValThrAasnProProThrAasnProProAlaGlnAaspSerHisPro 600
 QY 2179 GCAGTCATTTGGTAGCAACTGCTGCTTCTGTTTCAAAATTAGTGGGCTATCTTTTGTAG 2238
 Db 601 AlaIlelleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620

QY 2239 GATTTCGATGATACAGCTTTATGATAGGTATGATGGCTAGGTTCTTAATCAAAAAATCAAT 2298
 Db 621 AaspLeuAaspThrAlaTyrAaspAlaGlyTyrAaspTyrLeuGlySerAasnGlnLysIleAasp 640
 QY 2299 GTCTCTGAAATACAGTTAGGACTAAGCCCCCAGCTAAGTAAAGCCCATCAGATTTGACTCTA 2358
 Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAasnAlaProSerAaspLeuThrLeu 660
 QY 2359 GCGAATGAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTTCGCTGGGATCCCTAAT 2418
 Db 661 GlyAasnGlnMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrPheAaspProasn 680
 QY 2419 ACAGCAAAATAATCGTCTTATATCTCTGAAAGCTACATGAGCTAAAACTGGGTATATCCT 2478
 Db 681 ThrAlaAasnAangGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAasnPro 700
 QY 2479 GGGCTGAGGAGTAGCTTCTTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
 Db 701 GlyProGluArgValAlaSerLeuValProAasnSerLeuTyrGlySerIleLeuAaspIle 720
 QY 2539 CGATCTGCGCATTCAGCAATTCGAATTTCTCTATCATCACCGCATGCTTTAGTTCAGGATAT 2598
 Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAaspGlyArgSerTyrCysArgGlyLeu 740
 QY 2599 TGGTCTTCTGAGGTTTCGAATTTCTCTATCATCACCGCATGCTTTAGTTCAGGATAT 2658
 Db 741 TrrValSerGlyValSerAasnPhePheTyrHisAaspArgAaspAlaLeuGlyGlnGlyTyr 760
 QY 2659 CGGTATATTAGTGGGGTTTATCTTAGGAGCAACTCTCTACTTTTGATCATCATGCTTT 2718
 Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAasnSerTyrPheGlySerSerMetPhe 780
 QY 2719 GGTCTACCATTTCCGAAGTATTGTTAGTAGATCTAAAGATTATGATGATGCTGCTTCCAAT 2778
 Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAaspTyrValValCysArgSerAasn 800
 QY 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCAACCAAGCTTTTATGTTGATTCCTAT 2838
 Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
 QY 2839 TTGTTTCGAGATGCTTTATCCGTCGCTAGCTACGGTTTGGGAATCAAGTATCAAGAAC 2898
 Db 821 LeuPheGlyAaspAlaPheLeuArgAlaSerTyrGlyPheGlyAasnGlnHisMetLysThr 840
 QY 2899 TCATATACATTTGCAGAGGAGCGATGTTCTGTTGGGATAATAACTGCTCTGCTGGAGAG 2958
 Db 841 SerTyrThrPheAlaGluLysAaspValArgTyrPheAasnAasnCysLeuValGlyGlu 860
 QY 2959 ATTGAGCGGATTAACGATTTGATTACTCCATCTAAGCTCTATTTGAATGAGTTGCGT 3018
 Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAasnGluLeuArg 880
 QY 3019 CCTTTCGTCGAACTGAGTTTCTTATCCCATCATGAATCTTTTACAGAGGAAGCGCAT 3078
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAaspHisGluSerPheThrGluGluGlyAasp 900
 QY 3079 CAAGCTCCGGCATTCAGAGCGGACATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138
 Db 901 GlnAlaAlaGalaPheArgSerGlyHisLeuMetAasnLeuSerValProValGlyValLys 920
 QY 3139 TTTGATTCAGATGTTCTTAGTACACATCTTAATAATATAGCTTTATCGCGCTTATCTGT 3198
 Db 921 PheAaspArgCysSerSerThrHisProAasnLysTyrSerPheMetGlyAlaTyrIleCys 940
 QY 3199 GATGCTTATCGCACCATCTCTGGTACTAGACAAAGCTCTCTATCCATCCATCAAGACATGG 3258
 Db 941 AaspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGlnThrTyr 960
 QY 3259 ACACAGAGTCCCTTTCATTTAGCAAGATGAGTGTGTGTGTAGAGGATCTATGTATGCT 3318
 Db 961 ThrThrAaspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980

QY 3319 TCTTAACAGTAATATAGAGTATATGCGCATGCGAGATAGTATGAGATGCTTCT 3378
DB 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
DB 1001 ArgGlyTyrGlyLeuSerAlaGlySerIleValArgPhe 1013

RESULT 18
ID AEA19010
XX AEA19010 standard; protein; 1013 AA.
AC AEA19010;
XX
XX
DT 28-JUL-2005 (first entry)
XX
DE Chlamydia trachomatis protein - SEQ ID 13.
XX
XX chlamydia trachomatis infection; antibacterial; vaccine.
KW
XX
OS Chlamydia trachomatis.
XX
XX
PN US2005106162-A1.
XX
XX
PD 19-MAY-2005.
XX
XX 16-DEC-2004; 2004US-00498327.
XX
XX 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002EP-00182233.
PR 14-AUG-2002; 2001EP-00218924.
PR 12-DEC-2002; 2002WO-1B005761.
XX
XX (GRAN/) GRANDI G.
PA (RATT/) RATTI G.
XX
XX Grandi G, Ratti G;
XX
XX WPI; 2005-354777/36.
DR N-PSDB; AEA19011.
XX
XX New immunogenic composition comprising Chlamydia trachomatis protein or
PT its encoding nucleic acid, useful for immunizing against, or treating or
PT preventing chlamydial infection.
PT
XX
XX Claim 13; SEQ ID NO 13; 90pp; English.
XX
XX The invention comprises an immunogenic composition for the treatment or
CC prevention of a Chlamydia trachomatis infection. The immunogenic
CC composition contains a Chlamydia trachomatis protein or its encoding
CC nucleic acid. The immunogenic composition of the invention is useful
CC treating or preventing chlamydial infection. The present amino acid
CC sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was
CC obtained from the USPTO website.
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX
XX Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5123.50 Matches: 979
Percent Similarity: 98.8% Conservative: 22
Best Local Similarity: 96.6% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 9 Gaps: 1

US-10-701-844-1 (1-4435) x AEA19010 (1-1013)

QY 382 ATGCAACGCTTTTCCATAAGTCTTCTTCTTCAATGATTTCTAGCTTATTTCTGCTCT 441
DB 1 MetGlnThrSerPheHisIleValPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20

QY 442 TTAATGGGGGGGATATGCGACGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
DB 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGAGGAGAGTAAACATTTAAATAATCTTGACAAATCTATTGCGAGCTTTGCTTTAACT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80
QY 622 TGTGTTGGGAACCTTATTAGGGAGTTTTTACTGTTTGGGAGAGGACACTCGTTGACTTTC 691
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTCTGATGACACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTCTTTCCAAATTGCAATTCATTACTTGCCTGA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACGCTCT 861
DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTTAAACAGATCTTTGTTTACTCAATATGAGAAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTCGAGATGGGAGCTTATAGATGCTAAGAGCTTAAACGGTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAAATTAGCAAGCTTTTGTCTCTTCCAGAAATAATCTGCTCAAGCTGATGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTACCGAGTTTCTGCTATGGCTTAACGAGGCTCTATTGCTTTGATGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
QY 1102 GTTGCAGGAGTAGAGGGGAGGATTGCTGCTGTTTCAGATGCGGACGAGGAGTGTCA 1161
DB 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAAGATCCAGTAGTAAGTTTTTCCAGAAATACTGCGGTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGAAAGTACCGGAGTAGAGGAGGATTTACTCTTACCGGAAACGTTGCTTTCCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAACCACTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu 320
QY 1342 CAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATTTACGAGATGGAGGAGCTATC 1401
DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTAAGAATCGTGGCAA---GCAGGATCCAAATCACTCTGGATCAGTTTCTTTGAT 1458
DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGGAGGAGTAGTTTCTTCTTACTAGCAATGTAGCTGCTGGAAAGGGGAGCTATTAT 1518
DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
QY 1519 GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAATTTTTTAAGGAATATCGCTAAT 1578

Db 381 AlalysylsLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
Qy 1579 QATGTCGAGCGATTTATTAGGAAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT 1698
Db 421 AsplleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
Qy 1699 GCGTAACTGTGCTCTCAAGCCATTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrLeu 460
Qy 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTAAATGATCCCATCGAGATGGCAACGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
Qy 1819 AACGACGAGCGCATCTTCCAACTTCTAAATAATTAACGATGGTGAAGGATACACAGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
Qy 1879 GATATTGTTTTGCTAAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGGA 1938
Db 501 AsplleIlePheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Qy 1939 AGGATTGTTCTTCGTAAGGCAAAATTTATCAGTGAATTCCTAAAGTCAGACAGGTGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Qy 1999 AGTCGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCCAACACCACCAACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
Qy 2059 CCTCTGCGCTAAATCAGTTGATCAGCTTCCCAATCTGCAATTCCTCTCTCTCTCTG 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
Qy 2119 TTAGCAAAACAATGCAAGTATGATAGGTATGATGCTAGGTCTTAATCAAAAAATCAAT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
Qy 2179 GAGTCATTTGTAGACAACTGCTGCTGTTCTGTACAAATTTAGTGGGCTATCTTTTGTAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy 2239 GATTTGGATGATACAGCTTATGATAGGTATGATGCTAGGTCTTAATCAAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
Qy 2299 GTTCGAAATTAACAGTTAGGACTAAGCCCCCAGCTAATGCCCCATCAGATTGTACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Qy 2359 GGGATAGATGCTTAAGTATGGCTATCAAGAGCTGGAAGCTTCCGTGGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
Qy 2419 ACAGCAAAATAATGGTCTTATCTGAAAGCTACATGACTAAAGCTGGGTATATCTCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
Qy 2479 GGGCTGAGCGAGTACTCTTTGTTCCAAATAGTTTATGGGATGCCATTTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
Qy 2539 CGATCTGCCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy 2599 TGGGTTCTCGAGTTTCGAATTTCTCTATCATGACCGGATGCTTTAGTCTCAGGATAT 2658

Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGGACAAATCTCTACTTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTGCTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCAACAAAGCTTTATTTGGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTTCGAGATGCTGTTTATCCGTCTAGCTACCGGTTTGGGAATCAGCATATGAAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2899 TCATATACATTTTCGAGAGGAGCGATGCTTCGTTGGGATAATACTGCTCGCTCGAGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGAGCGGATTTACCGATTGTGATTACTCCATCTAAAGCTCTATTGAAATGAGTTGCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy 3019 CCTTTCGTGGAAGCTGAGTTTCTTATGCCGATCATGAAATCTTTTACAGAGGAAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGCATTTCAAGAGCGGACATCTCTAAATCTATCATGTTCCCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCATGTTCTAGTACATCCTCTAATAATAATAGCTTTTATGGCGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCGCACCATCTCTGGTACTGAGCAACAGCTCCTATCCCATCAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATCGAGTTGTGGTTAGAGGATCTATGTATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetLysAla 980
Qy 3319 TCTCTAAACAAGTAATATAGAAATATATGCGCATGGAAGATATGAGTATCGAGATCCTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 19

AEAL9000
ID AEAL9000 standard; protein; 1013 AA.
AC AEAL9000;
XX
XX 28-JUL-2005 (first entry)
DT
XX Chlamydia trachomatis protein - SEQ ID 3.
DE Chlamydia trachomatis infection; antibacterial; vaccine.
KW Chlamydia trachomatis.
OS Chlamydia trachomatis.
XX
PN US2005106162-A1.
XX
PD 19-MAY-2005.
XX
PP 16-DEC-2004; 2004US-00498327.

XX 12-DEC-2001; 2001GB-00029732.
 PR 06-AUG-2002; 2002EP-00182233.
 PR 14-AUG-2002; 2001EP-00218924.
 PR 12-DEC-2002; 2002WO-18005761.
 XX (GRAN/) GRANDI G.
 PA (RATT/) RATTI G.
 XX Grandi G, Ratti G;
 XX WPI; 2005-354777/36.
 DR N-PSDB; AEA19001.
 XX New immunogenic composition comprising Chlamydia trachomatis protein or
 PT its encoding nucleic acid, useful for immunizing against, or treating or
 PT preventing chlamydial infection.
 XX Claim 13; SEQ ID NO 3; 90pp; English.
 XX The invention comprises an immunogenic composition for the treatment or
 CC prevention of a Chlamydia trachomatis infection. The immunogenic
 CC composition contains a Chlamydia trachomatis protein or its encoding
 CC nucleic acid. The immunogenic composition of the invention is useful for
 CC treating or preventing chlamydial infection. The present amino acid
 CC sequence represents a Chlamydia trachomatis protein of the invention.
 CC NOTE: The present sequence is not shown in the specification, but was
 CC obtained from the USPTO website -
 CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
 XX Sequence 1013 AA;
 SQ

Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5121.50 Matches: 978
 Percent Similarity: 98.8% Conservative: 23
 Best Local Similarity: 96.5% Mismatches: 11
 Query Match: 65.0% Indels: 1
 DB: 9 Gaps: 1

US-10-701-844-1 (1-4435) x AEA19000 (1-1013)

QY 382 ATGCAACGCTTTCCATAAGTCTTCTTCAATGATTTAGCTTATTTCTGCTCT 441
 DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetLeuLeuAlaTyrSerCysSer 20
 QY 442 TTAATGGGGGGGATATGACAGCAAAATCATGCTTCTCAAGAAATTTACGATGGGAG 501
 DB 21 LeuSerGlyGlyTyrAlaAlaGluLeuMetIleProGlnGlyIleTyrAspGlyGlu 40
 QY 502 ACCTTAATCTATCTTCCATATCTGTTATAGGAGATCCGATGGGACTCTGTTTTT 561
 DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 QY 562 TCTGCAGGAGTAAATTAATAAATCTTGACAAATCTATTGCGAGTTTGCCTTTAAGT 621
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80
 QY 622 TGTTCGGGAACTATTAGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 QY 682 GAGAACATACGACTTCTACAAATGGGGAGCTCTAAGTAATAGCGTCTGATGGGACTG 741
 DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 QY 742 TTTACTATTGAGGGTTTAAAGAAATTAATCTTTTCCAAATTCGCAATTCATTCTGCGCTA 801
 DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
 QY 802 CTCGCTCTCCACGACTAATAGGCTAGCCAGACTCCGACGACAAATCTACACCGTCT 861
 DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160

QY 862 AATGGTACTATTATTCTAAACACAGATCTTTTGTTACTCAATAATGAGAAGTTCTCATTC 921
 DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
 QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
 DB 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
 QY 982 CGAATTACGACGCTTTGTCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTCT 1041
 DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys 220
 QY 1042 CAAGTAGTCACCAAGTTTCTCTCTATGCTTAACGAGGCTCTTATTCCTTTGTCGCAAT 1101
 DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
 QY 1102 GTTCAGGAGTAAGAGGGGAGGATTTGCTGCTGTTCCAGATGGGAGGAGGAGTCTCA 1161
 DB 241 ValAlaGlyValArgGlyGlyLeuAlaAlaValGlnAspGlyGlnGlyValSer 260
 QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGGTAGAGTTT 1221
 DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 QY 1222 GATGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACTGTTGTTCTCTG 1281
 DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
 QY 1282 AATAATGGAAAAACCTTTGTTCTCAACAAATGTTGCTTCTCTGTTTACATTTGCTGAAG 1341
 DB 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrLeuAlaAlaGlu 320
 QY 1342 CAACCAACAGTGGAGCAGCTTCTAATACAGTAATAATACGAGATGGAGGAGCTATC 1401
 DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340
 QY 1402 TTCTGTAAAGAAATGGTGGCAA--GCAGGATCCAAATAACTCTCGATCAGTTTCTCTTGAT 1458
 DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 QY 1459 GGAGAGGGAGTAGTTTCTTTTAGTAGCAATGTAGTCTGGGAAAGGGGAGCTATTTAT 1518
 DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380
 QY 1519 GCCAAAGCTCTCGGTTGCTAATCTGTCCTGTAACAATTTTAAAGATATCGCTAAT 1578
 DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 QY 1579 GATGGTGGAGCGATTATTTAGGAGATCTGGAGAGCTCAGATTTATCTGCTGATTTATGGA 1638
 DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
 QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACACAGCCAAAGAGATGTCGCCGATGTTAAT 1698
 DB 421 AspiIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
 QY 1699 GCGGTAACTGTGTCCTCAAGCCATTTGATGGGATCGGGAGGAGGAGGAGGAGGAGT 1758
 DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
 QY 1759 AGAGCTTAAGCAGCGGATCAGATTTCTTTTAAATGATCCCATCCAGATGGCAACGGAAAT 1818
 DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
 QY 1819 AACCAGCCAGCGAGCTCTCCAACTTCTAAAAAATAACCATGCTGAGGATACACAGGG 1878
 DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyrThrGly 500
 QY 1879 GATATTGTTTCTAATGGAACGAGTCTTTGTATCCAAATGTTACGATAGGAGGAGGAG 1938
 DB 501 AspiIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520

QY 1939 AGGATTGTTCTTCGTAAGGCAAAATTTATCAGTGAATTTCTTAAGTCACAGAGGTGG 1998
DB ArgilleValLeuArgGluValAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAGCTGGAGTACATGGGATTTTGTAACTCCAAACCCACCAACAG 2058
DB SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
QY 2059 CCTCTCGCGGTAAATCAGTTGATCAGCTTTCCAAATCTGCATTTGTCTCTTCTTCTTTG 2118
DB ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAACAATGCAAGTTAGCAATCCTCTCAATCTCCAGCCCAAGATTTCTCATCCT 2178
DB LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCATTGGTAGCACAACTGCTGGTCTGTGTACAAATAGTAGGCGCTATCTTTTTGAG 2238
DB AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298
DB AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTCTGAATTTACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358
DB ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGAATGAGATGCTCAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCGGATCCTCAAT 2418
DB GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680
QY 2419 ACAGCAAAATAATGGTCTTATCTCTGAAAGTACATGAGCTAAACCTGGGTAAATCCT 2478
DB ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCTGAGCAGTAGCTCTTTGGTTCCTCAATAGTTTATGGGATCCATTTTACATATA 2538
DB GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGTCGAGATTA 2598
DB ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTCTGGAGTTTCGAATTTCTTATCATGACCCGATGCTTTAGTTCAGGATAT 2658
DB TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTACTTTTGGATCATCGATGTTT 2718
DB ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGTAGATCTTAAAGATTATGTAGTGTGCTTCCAAT 2778
DB GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTTATGTGCATCTAT 2838
DB HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTCCGGAGATCGGTTTATTCGTCTAGCTAGCTAGGTTTGGGAATCAGCATATGAACCC 2898
DB LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGAGCGATGTTCTGTGGATAATAACTGTCTGCTGGAGAG 2958
DB SerTyrThrPheAlaGluGluSerAspValAlaGlyTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGGACGGGATTACCGATTGTGATTACTCCATCTTAAGCTCTTATTGAATGAGTTGCT 3018
DB IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CCTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGAT 3078

DB PropheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCGATTCACAGAGCGGACATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138
DB GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTGTACACATCCCTAAATAATATAGCTTTATGGCGCTTATATCTGT 3198
DB PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAAGCTCTCTATCCATCAAGAGACATGG 3258
DB AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTCTTCAATTAGCAAGACATGGAGTTGTGGTAGAGGATCTATGATGCT 3318
DB ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTTAAACAATAATAGATATATGGCCATCGACATATGAGTATCGAGATGCTTCT 3378
DB SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGCTATGCTTTGAGTCAGGAAGTAGAGTCCGGTTC 3417
DB ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 20
AEA18998
ID AEA18998 standard; protein; 1013 AA.
XX AEA18998;
XX 28-JUL-2005 (first entry)
XX Chlamydia trachomatis protein - SEQ ID 1.
XX chlamydia trachomatis infection; antibacterial; vaccine.
XX Chlamydia trachomatis.
XX US2005106162-A1.
XX 19-MAY-2005.
XX 16-DEC-2004; 2004US-00498327.
XX 12-DEC-2001; 2001GB-00029732.
XX 06-AUG-2002; 2002EP-00182233.
XX 14-AUG-2002; 2001EP-00218924.
XX 12-DEC-2002; 2002WO-IB005761.
XX (GRAN/) GRANDI G.
XX (RATT/) RATTI G.
XX Grandi G, Ratti G;
XX WPI; 2005-354777/36.
XX N-PSDB; AEA18999.
XX New immunogenic composition comprising Chlamydia trachomatis protein or its encoding nucleic acid, useful for immunizing against, or treating or preventing chlamydial infection.
XX Claim 13; SEQ ID NO 1; 90pp; English.
XX The invention comprises an immunogenic composition for the treatment or prevention of a Chlamydia trachomatis infection. The immunogenic composition contains a Chlamydia trachomatis protein or its encoding nucleic acid. The immunogenic composition of the invention is useful for treating or preventing chlamydial infection. The present amino acid sequence represents a Chlamydia trachomatis protein of the invention.
XX NOTE: The present sequence is not shown in the specification, but was

CC obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX
SQ Sequence 1013 AA;

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5121.50 Matches: 978
Percent Similarity: 98.8% Conservative: 23
Best Local Similarity: 96.5% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 9 Gaps: 1

US-10-701-844-1 (1-4435) x AEA18998 (1-1013)

```
QY 382 ATGCAACGCTCTTCCATAGTTCCTTCTTCAATGATCTAGCTTATCTGCTGCTCT 441
DB 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuLeuAlaTyrSerCysCysSer 20
QY 442 TTAATGGGGGGATATGCGAGCAAAATCATGTTCTCCACAGGAATTTACGATGGGAG 501
DB 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACCTTAACGTATCATTTCCCTACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGATTAACTTAAAAATCTTGACAATCTTATTCAGCTTTCCTTTAAAGT 621
DB 61 SerAlaGlyLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAACTTATAGGAGTTTACTGTTTTTAGGGAGAGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGAGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAATTATCTCTTTTCCAAATTCGAATTCATTACTTGGCGTA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTCCTGCTGCACGACTAATAAGGTAGCGAGCTCCGAGCTCCGACGACCAATCTACCGTCT 861
DB 141 LeuProAlaAlaThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
QY 862 AATGCTACTATTATTCTTAAACAGATCTTTGTTACTCAATATGAGAGTCTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuLeuAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGAGCTTTAAGCTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTGTCTTCCAGAAAAATCTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTACTCACCAGTTTCTGCTATGCTTAAACGAGGCTCTATGCTTGTGCGCAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProfileAlaPheIleAlaAsn 240
QY 1102 GTTGCAGGATTAAGAGGGGAGGATTTGCTGCTCTTCCAGATGGGACGAGGAGTGTCA 1161
DB 241 ValAlaGlyValArgGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATATCTCGGTTAGAGTTT 1221
DB 261 SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTCCTACGGGACGTGCTTCTG 1281
```

```
DB 281 AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAACAACTTGTCTTCTCACAATGTTGCTTCTCCGTTTACATTCGCTGAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu 320
QY 1342 CAACCAACAAGTGGAGCAGGCTTCTAATACGAGTAATAATATTCGAGAGATCGAGAGCTATC 1401
DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTAGAAGTGGTGGCAA---GCAGGATCCATACTCGATCGATTCAGTTCTTCTTGTAT 1458
DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 CGAGGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGACTATTTAT 1518
DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380
QY 1519 GCCAAAAAGCTCTCGGTTCTAATCTGTGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATGTGGAGCGATTTATTTAGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATATGGA 1638
DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCAAGAGAGATCTGCGGATGTTAAT 1698
DB 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGCTAATCTGTCTCACAGCCATTTCCATGGGATCGGAGGGGAAAAATAACGACATTA 1758
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGCATCTCTTTAATGATCCCATGAGATGGCAAGCGGAAAT 1818
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACCAGCCAGCGAGTCTTCCAACTTCTAATAATTAACGATGTTGAGGATACACAGGG 1878
DB 481 AsnGlnProAlaGlnSerGluProLeuLysLysAlaAsnAspGlyGluGlyTyrThrGly 500
QY 1879 GATATTGTTTCTAATGAGACGATCTCTTTCTACCAAAATGTTACGATAGACGAGGA 1938
DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTGTTCTTCGTGAAAGGCCAAAATATATCAGTGAATCTCTAAAGTCAGAGGTGGG 1998
DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAAACCACACAG 2058
DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCAATTTGTCTTTCTTCTTCTT 2118
DB 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAAAACAATGCAGTTACGAATCTCTCTACCAATCTCTCCAGCGCAAGATCTCATCCT 2178
DB 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCAATTGGTAGCAACTGCTGTTCTGTTTACAAATTAGTGGGCTATCTTTTGTAG 2238
DB 601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProfilePhePheGlu 620
QY 2239 GATTTGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298
DB 621 AspLeuAspAspThrAlaTyrAspArgTyrAspThrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTGTAATTTACAGTTAGGAGCTTAAGCCCCCAGCTAATCCCCCATCAGATTTGACTTA 2358
DB 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
```



```
QY 2359 GGGAAATGAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGATCCTAAT 2418
Db 661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY 2419 ACAGCAATATGGTCCCTTATCTACTGTAAGACTACATGAGCTAAAGCTGGGTATAATCCT 2478
Db 681 ThrAlaAenAenglyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro 700
QY 2479 GGGCTGAGGAGTACCTTCTTTGGTCCAAATAGTTTATGGGATCCATTTTATGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGCGCAATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATGTGCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTCTATCATGACCGGATGCTTATAGGTCAGGGATAT 2658
Db 741 TrpValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTATTCTCTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCAATTCAGCAATTTGGTAGATCTAAAGATTATGTAGTGTGTGTTCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
QY 2779 CATCATGCTTGATAGGATCGGTTTATCTATCTACCAACAGCTTATGTGATCCAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTTCGAGAGTCCGTTTCTCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGCGATGTTCTGTTGGATATAATCTGCTGCTGGAGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValArgTrpAspAenAenCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGATTCAGGATTCGATTACTCTTACCTTACCTTATTTGATGAGTGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
QY 3019 CTTTTCGTCGACGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGACCGAT 3078
Db 881 PropheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCGATTCAAGAGCGGACATCTCTAAATCTATCATAGTTCTGCTGGAGTGA 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACATCTCTAATAATATAGCTTTATGCGGCTTATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGCTACTGAGACAACGCTCTCTATCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGACAGACATGAGGATGTTGGTTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTTACAGATTAATAGAGTATATGCCATGGAAGTATGAGTATGAGTATGAGTCTCT 3378
Db 981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

RESULT 21

```
AEA19014
ID AEA19014 standard; protein; 1013 AA.
XX AC AEA19014;
XX DT 28-JUL-2005 (first entry)
XX DE Chlamydia trachomatis protein - SEQ ID 17.
XX KW chlamydia trachomatis infection; antibacterial; vaccine.
XX OS Chlamydia trachomatis.
XX PN US2005106162-A1.
XX PD 19-MAY-2005.
XX PF 16-DEC-2004; 2004US-00498327.
XX PR 12-DEC-2001; 2001GB-00029732.
XX PR 06-AUG-2002; 2002EP-00182233.
XX PR 14-AUG-2002; 2001EP-00218924.
XX PR 12-DEC-2002; 2002WO-IB005761.
XX (GRAN/) GRANDI G.
XX PA (RATT/) RATTI G.
XX PI Grandi G, Ratti G;
XX DR WPI; 2005-354777/36.
XX DR N-PDDB; AEA19015.
XX PT New immunogenic composition comprising Chlamydia trachomatis protein or
PT its encoding nucleic acid, useful for immunizing against, or treating or
PT preventing chlamydial infection.
XX PS Claim 13; SEQ ID NO 17; 90pp; English.
XX CC The invention comprises an immunogenic composition for the treatment or
XX prevention of a Chlamydia trachomatis infection. The immunogenic
XX composition contains a Chlamydia trachomatis protein or its encoding
XX nucleic acid. The immunogenic composition of the invention is useful for
XX treating or preventing chlamydial infection. The present amino acid
XX sequence represents a Chlamydia trachomatis protein of the invention.
XX NOTE: The present sequence is not shown in the specification, but was
XX obtained from the USPTO website -
XX seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX SQ Sequence 1013 AA;
```

Alignment Scores:

Pred. No.:	0	Length:	1013
Score:	5112.50	Matches:	977
Percent Similarity:	98.7%	Conservative:	23
Best Local Similarity:	96.4%	Mismatches:	12
Query Match:	64.9%	Indels:	1
DB:	9	Gaps:	1

US-10-701-844-1 (1-4435) x AEA19014 (1-1013)

```
QY 382 ATGCAAGCTCTTTTCATAAGTCTTTCTTCAATGATCTAGCTATTCTTCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGGATATGCACAGAAATCATGTTTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACCTTAAGTGTATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
```

Qy	562	TCTGCAGGAGATTAA	CAATTA	AAAAATCTTTGACA	ATTCTTA	TTCAGC	TTTGGCTTTAA	GTT	621
Db	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer							80
Qy	622	TGTTTTGGGAACCTTAT	TAGGAGCTTTTAC	TGTTTTTAGGGAGGAGGAC	CAC	TGTTGAC	TTTC		681
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe							100
Qy	682	GAGAAACATACGACTTCTA	CAAAATGGGGCAGCTCTA	AGTAATTAACCGCTCTG	TGATG	ACTG			741
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaLeuSerAspSerAlaAsnSerGlyLeu							120
Qy	742	TTTACTATTGAGGGTTTTAA	GAATTTATCTCTTTTCCA	ATTGCAATTCATTACT	TTGCGCGTA				801
Db	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal							140
Qy	802	CTGCCTGCTCCACGACTA	TAAAGGTAGCCAGACTCC	GACGACCAACATCTAC	ACCGTCT				861
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer							160
Qy	862	AATGGTACTATTATTCTTAA	AAACAGACTTTTGTGTACT	CAATTAATGAGA	GTTCCTCATTC				921
Db	161	AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe							180
Qy	922	TATAGTAAATTTAGTCTCT	GAGATGGGGAGCTATAG	ATCCTAAGAGCTTAA	CGTTCAA				981
Db	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln							200
Qy	982	GGAATTAGCAAGCTTTGTG	CTCTCCAGAAAATACTGCT	CAAGCTGATGGGGAGCTGT					1041
Db	201	GlyIleSerLysLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys							220
Qy	1042	CAAGTAGTCCACAGTTTCT	CTGCTATCGCTTAA	CGAGGCTCTATTGCTTTG	TAGCGGAAT				1101
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn							240
Qy	1102	GTTGCAGGAGTAAGAGGG	GAGGATGTCGCTGTT	CAGATGGCGACGAGGAGTG	TCA				1161
Db	241	ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer							260
Qy	1162	TCATCTACTTCAACAGA	AGATCCAGTAGTAGATT	TTTCCAGAAATAC	TGCGGTAGAGTTT				1221
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe							280
Qy	1222	GATCGGAACGTAGCCGAG	TAGGAGGAGATTTACT	CTACGGGAACGTTGCTTCT	CTCGT				1281
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu							300
Qy	1282	AATAATGGAAAAACCTT	GTTTTCTCAACAATGTG	CTCTCTCTGTTTACAT	TGCTGCTAAG				1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu							320
Qy	1342	CAACCAACAGTGGACAG	GGCTTCTAATACGAGTA	ATAATTACGGAGATGGAGGAGCTATC					1401
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle							340
Qy	1402	TTCTGTGAAGATGTGCG	CCAA--GCGAGATCCAA	TAACTCTGATCAG	TTTTCTCTTTGAT				1458
Db	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp							360
Qy	1459	CGAGAGGAGTAGTTTTCT	TTTAGTACCAATAGCT	CTGCTGGAAAGGGGAGCTATTTAT					1518
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTy							380
Qy	1519	GCCAAAAGCTCTCGTT	GTCTAACTGTGGCCCTG	ACAAATTTTTTAAAGGAATAT	CGCTTAAT				1578
Db	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn							400
Qy	1579	GATCGTGGAGCGATTTAT	TTTAGGAGAATCTG	CGAGGCTCAGTTTATCT	GTGATTATGGA				1638
Db	401	AspGlyGlyAlaIleTyLeuGlyGlnSerGlyGlnLeuSerLeuSerAlaAspTyGly							420
Qy	1639	GATATTATTTTCGATGG	GAATCTTTAAAGAA	CACGCCAAGAGAAT	GTGCGCGATGTTAAT				1698

421 AspllellePheAspGlyAsnLeuIysArgThrAlaIysGluAsnAlaIaAspValAsn 440

1699 GGCCTAACTGTCCTCCACAAAGCCATTTTCGATGGGATCGGAGGGAATAACGACATTA 1758

441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyPheIysIleThrThrLeu 460

1759 AGAGCTAAACAGGGCATCAGATTCTCTTTAATGATCCATCGAGATGGCAACGGAAAT 1818

461 ArgAlaIysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480

1819 AACCAAGCCAGCGAGTCTCCAAACTTCATAAAATTAACGATGGTGAAGGATACACAGGG 1878

481 AsnGlnProAlaGlnSerGluProLeuIysIleAsnAspGlyGluGlyTyrThrGly 500

1879 GATATTCTTTTGTCTAATGGAAGCAGTACTTTGTATCCAAAATTTGATAGACAAAGGA 1938

501 AsplleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520

1939 AGGATTGCTTCTGTAAGGAAGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998

521 ArgIleValLeuArgGluIysAlaIysLeuSerValAsnSerLeuSerGlnThrGlyGly 540

1999 AGTCGTATATGGAAGCTGGAGTACATCGGATTTGTAACTCCACAACACCAACAAACAG 2058

541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560

2059 CCTCTCCGCGTAATCAGTTGATCAGCGCTTCCCAATCTGCATTTGTCTCTTTCTCTTGTG 2118

561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580

2119 TTAGCAACAATGAGTTAGCAATCTCTCTCAATCTCCAGCGCAAGATCTCATCCT 2178

581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600

2179 GCAGTCAATGGTAGCAAACTGCTGGTCTGTGTACAAATTAGTGGGCCCTATCTTTTGTGAG 2238

601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620

2239 GATTGTGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAATCAAAAAATCAAT 2298

621 AsplleuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnIysAlaAsp 640

2299 GTCTCGAAATPACAGTTAGGGAATAAGCCCGCAGCTAATGCCCATCAGATTGACTCTTA 2358

641 ValLeuIysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660

2359 GGGATGAGATGCTAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGGATCTTAAT 2418

661 GlyAsnGluMetProIysTyrGlyTyrGlnIysSerTrpIysLeuAlaTrpAspProAsn 680

2419 ACAGCAAAATATGCTCTTATATCTCGAAAGCTACATGGAATCAAAAACCTGGGTATAATCCT 2478

681 ThrAlaAsnAsnGlyProTyrThrLeuIysAlaThrTrpThrLysThrGlyTyrAsnPro 700

2479 GGGCTGACGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATTCGATTTAGATATA 2538

701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720

2539 CGATCTGGCAATTCAGCAATTCAGCAAGTGGATGGCGGCTCTATTGTCGAGGATTA 2598

721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740

2599 TGGGTTTCTGAGTTTCGAATTTCTTCTATCATGACCGGATGCTTTAGTCAGGGATAT 2658

741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760

2659 CGGTATATTAGTGGGGTTATCTCTTAGGAGCAAACTCTCACTTTGGATCATCGATGTTT 2718

761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780

2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTTAAGATATGCTAGTGTCGTCTCCAAAT 2778

Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
 Qy 2779 CATCATGCTGCATAGAGTCCGTTTATCTATCTATCCCAACAGCTTTTATGTGATCCTAT 2838
 Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
 Qy 2839 TTGTTGAGAGTGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAACACC 2898
 Db 821 LeupheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
 Qy 2899 TCATATACATTCGAGAGAGCGCATCTTGGTGGGATATACATCTGCTGGCTGGAG 2958
 Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
 Qy 2959 ATTGAGCGGATACCCATTTGTGATTAATCTCATCTAAGCTCTATTTCAATGAGTTGGT 3018
 Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
 Qy 3019 CTTTTCGTCGAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGAGCGCAT 3078
 Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
 Qy 3079 CAAGCTCGGCATTCACAGAGCGGCATCTCTAAATCTATCAGTTCCCTGTTGGAGTGAAG 3138
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
 Qy 3139 TTTGATCATGTTCTAGTACACATCCTTAATAAATAGCTTTTATGGCGCTTATATCTGT 3198
 Db 921 PheAspArgCysSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
 Qy 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACACGCTCTATCCATCAAGAGACATGG 3258
 Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
 Qy 3259 ACAACAGATGCTTTTCATTACACAGACATGGAGTTGTGTTAGAGGATCTATGATGCT 3318
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 Qy 3319 TCTCTAACAAAGTAATAGAGTATATGCCATGGAAGATATGATGATCGAGATGCTTCT 3378
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 Qy 3379 CGAGCTAGTTTGAAGTCAGAGAAAGTAGATCCGGTTC 3417
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 22

AAV16737

ID AAV16737 standard; protein; 1013 AA.

XX AC AAV16737;

XX DT 21-JUL-1999 (first entry)

XX DE C. trachomatis B serovar HMW protein.

XX KW Chlamydia; high molecular weight protein; HMW protein; urethritis;

XX KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;

XX KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;

XX KW salpingitis; tubal occlusion; infertility; cervical cancer;

XX KW arteriosclerosis; atherosclerosis.

XX OS Chlamydia trachomatis.

XX PN W09917741-AL.

XX PD 15-APR-1999.

XX PF 01-OCT-1998; 98MO-US020737.

XX PR 02-OCT-1997; 97US-00942596.

XX XX (ANTE-) ANTEX BIOLOGICS INC.

XX FI

Jackson JW, Pace JL;

XX DR

WPI; 1999-287659/24.

XX XT

New Chlamydia protein useful for treating conjunctivitis, urethritis and cervical cancer.

XX PT

XX PS

Claim 4; Page 115-119; 141pp; English.

XX CC

The invention relates to an isolated Chlamydia species high molecular weight (HMW) protein having an apparent mol. wt. of 105-115 kD as determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can be used for preventing, treating or ameliorating a disorder related to Chlamydia e.g. bacterial infection, conjunctivitis, urethritis, lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis, pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical cancer, infertility, arteriosclerosis and atherosclerosis. The products can also be used for detection and diagnosis. The present sequence represents a C. trachomatis HMW protein

SQ Sequence 1013 AA;

Alignment Scores:

Pred. No.: 0 Length: 1013
 Score: 5106.50 Matches: 980
 Percent Similarity: 98.4% Conservative: 17
 Best Local Similarity: 96.7% Mismatches: 15
 Query Match: 64.8% Indels: 1
 DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x AAY16737 (1-1013)

Qy 382 ATGCAAAAGCTCTTCCATAAGTCTCTTCTTCAATGATTCTAGCTTATTTGCTGCTCT 441

Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20

Qy 442 TTAATGGGGGGGATATGACAGAAATCATGTTCTCTCAAGGAATTTACATGGGAG 501

Db 21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40

Qy 502 ACGTTAAGTGTATCATTTCCCTATCTATAGAGATCCGAGTGGCAGTACTGTTTTT 561

Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60

Qy 562 TCTGCAGGAGAGTTAAACATTAAAAATCTTGACAAATCTTATTCAGCTTTTGCCTTAAAGT 621

Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80

Qy 622 TGTGTTGGAACTTATTAGGAGTTTACTGTGTTTAGGAGGAGGACACTCGTTGACTTTC 681

Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100

Qy 682 GAGACATACCGACTTCTACAANTGGGCGACTCTAAGTAATAGCGCTGCTGATGACTG 741

Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120

Qy 742 TTTACTATTGAGGGTTTTAAAGAATTATCTTTTCCAAATTCGAATTCATTCTTCGCGTA 801

Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140

Qy 802 CTGCTGCTGCAACGACTTAATAAGGGTAGCCAGACTCCGACGACAACTCTACCGTCT 861

Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrSerThrProSer 160

Qy 862 AATGGTACTATTATTCTTAAACAGACTCTTTGTTTACTCAATAATAGAGAGTTCTCATTC 921

Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180

Qy 922 TATAGTAATTTAGTCTCTGGAGATCGGGAGCTATAGATGCTTAAGAGCTTAAACGGTCAA 981

Db 181 TyrSerAsnSerValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200

QY 982 GGAATTACAGCTTTGCTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
DB |||||
QY 201 GlyIleSerLysLeuCyvalPheGlnGluAenThrAlaGlnAlaAspGlyAlaCys 220
DB |||||
QY 1042 CAAGTAGTCACCAAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCCCTTTAGCGAAT 1101
DB |||||
QY 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheValAlaAen 240
DB |||||
QY 1102 GTTCAGGAGTAGAAGGGGAGGATGCTGCTGCTTCAGGATGGCCAGGAGGAGTCA 1161
DB |||||
QY 241 ValAlaGlyValLargGlyGlyAlaAlaValGlnAenGlyGlnGlnGlyValSer 260
DB |||||
QY 1162 TCATCTACTTCAACAGACAGCCAGTAGTAAGTTTTCAGAAATACTCCGCTAGAGTTT 1221
DB |||||
QY 261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
DB |||||
QY 1222 GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACCTGCTTTCCTG 1281
DB |||||
QY 281 AspGlyAenValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAenValAlaPheLeu 300
DB |||||
QY 1282 AATAATGGAAAAACCTTGTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341
DB |||||
QY 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaGlu 320
DB |||||
QY 1342 CAACCAACAGTGGCAGGCTCTTAATACGAGTAATAATTACGAGATGGAGGAGCTATC 1401
DB |||||
QY 321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAenTyrGlyAspGlyAlaIle 340
DB |||||
QY 1402 TTCTGTAGAAGTGGTCGCGCAA---GCAGGATCCAATAACTCTGGATCAGTTTCTTTGAT 1458
DB |||||
QY 341 PheCyLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAsp 360
DB |||||
QY 1459 GGAGGAGGAGTAGTTTCTTTAGTACAAATGATGCTGCGGAAAGGGGAGCTATTAT 1518
DB |||||
QY 361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyAlaIleTyr 380
DB |||||
QY 1519 GCMAAAGCTCTCGGCTCTAAGTCTGCGCCGTACAAATTTTAAGGAATATCGCTAAT 1578
DB |||||
QY 381 AlaLysLysLeuSerValAlaAenCyGlyProValGlnLeuLeuGlyAenIleAlaAen 400
DB |||||
QY 1579 GATGTTGAGCGGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
DB |||||
QY 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
DB |||||
QY 1639 GATATTATTTTCAGTGGGAATCTTAAAGAACGCCAAGAGAGATGCTGCCGATGTTAAT 1698
DB |||||
QY 421 AspMetIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440
DB |||||
QY 1699 GCGGTAACTGTGCTCACAGCCATTTGATCGGATCGGAGGAGGAAATAACGACATTA 1758
DB |||||
QY 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIleThrThrLeu 460
DB |||||
QY 1759 AGAGCTAAAGCGGATCAGATTCTCTTAAATGATCCCATCGAGATGGCAACCGAAAT 1818
DB |||||
QY 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAspProIleGluMetAlaAenGlyAen 480
DB |||||
QY 1819 AACGAGCACGCGAGTCTTCCAAACTTCTAAATAATTAACGATGGTGAAGATACACAGG 1878
DB |||||
QY 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGluGlyTyrThrGly 500
DB |||||
QY 1879 GATATTGTTTTGCTAATGAGACAGTACTTTGTACCAAAATGTTACGATAGGACGAGGA 1938
DB |||||
QY 501 AspIleValPheAlaAenGlyAenSerThrLeuTyrGlnAenValThrIleGluGlnGly 520
DB |||||
QY 1939 AGGATTTCTTCTGTGAAAAGCGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
DB |||||
QY 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540
DB |||||
QY 1999 AGTCTGTATATGGAAGCTGGAGGATCATGGGATTTTGTAACTCCCAACCAACCAACAG 2058
DB |||||
QY 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
DB |||||
QY 2059 CCTCTCGCGCTAATCAGTTGATCACGCTTCCAAATCTGCAATTTGTTCTTCTTCTTGT 2118
DB |||||

DB |||||
QY 561 ProProAlaAlaAenGlnSerIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeu 580
DB |||||
QY 2119 TTGCAACAATGCAAGTTTACGAATCTCTACCAATCTCTCCAGCGCAAGATTCTCATCT 2178
DB |||||
QY 581 LeuAlaAenAenAlaValThrAenProThrAenProProAlaGlnAspSerHisPro 600
DB |||||
QY 2179 GCAGTCATTTGGTAGCACAACTGCTGTTTCTGTACAAATTAGTGGCCCTATCTTTTGTAG 2238
DB |||||
QY 601 AlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
DB |||||
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATTCGCTAGGTCTTCTTAATCAAAAATCAAT 2298
DB |||||
QY 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAenGlnLysIleAsp 640
DB |||||
QY 2299 GTTCTGAATTTACAGTTAGGAGTAAGCCCCAGCTAAATGCCCAATACAGATTGACTCTA 2358
DB |||||
QY 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAenAlaProSerAspLeuThrLeu 660
DB |||||
QY 2359 GCGAATGAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCTAAT 2418
DB |||||
QY 661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAen 680
DB |||||
QY 2419 ACAGCAAAATAATGGTCCCTTATCTCTGAAAGCTACATGGAATAAACTGGGTATATCT 2478
DB |||||
QY 681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAenPro 700
DB |||||
QY 2479 GGGCTGAGCGAGTAGTCTTCTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
DB |||||
QY 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTyrLysIleLeuAspIle 720
DB |||||
QY 2539 CCATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTGCGAGATTA 2598
DB |||||
QY 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCyArgGlyLeu 740
DB |||||
QY 2599 TGGGTTCTGGAGTTTCGAATTTCTTATCATATGACCCGATGCTTTAGTCTAGGATAT 2658
DB |||||
QY 741 TrpValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
DB |||||
QY 2659 CGGTATATTAGTGGGGTATTCTCTTAGGAGCAAACTCTTACTTTGATCATCTCGATGTTT 2718
DB |||||
QY 761 ArgTyrIleSerGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
DB |||||
QY 2719 GGTCTAGCATTTACCCAAAGTATTGTGTAGATCTAAAGATTTATGTAGTGTCTGCCAAT 2778
DB |||||
QY 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
DB |||||
QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTACCCAAAGCTTTTATGTGATCTTAT 2838
DB |||||
QY 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
DB |||||
QY 2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGTACGGGTTTGGGAATCAGCATATGAAAAACC 2898
DB |||||
QY 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
DB |||||
QY 2899 TCATATACATTTCCAGAGGAGGAGCGATGTTGCTGGGATAATACTGCTGGCTGGAGAG 2958
DB |||||
QY 841 SerTyrThrPheAlaGluLeuSerAspValCysTrpAspAenAenCysLeuValGlyGlu 860
DB |||||
QY 2959 ATTGGAGCGGGATTACCGATTGTGATTACTTCCATCTAAGCTCTATTGAAATGATTCGGT 3018
DB |||||
QY 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
DB |||||
QY 3019 CCTTTCGTGCAAGCTCAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGGCGAT 3078
DB |||||
QY 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
DB |||||
QY 3079 CAGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTTCTGTTGGAGTCAAG 3138
DB |||||
QY 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
DB |||||
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATATAATATAGCTTTATGGCGGCTTATCTGT 3198
DB |||||

Db 921 PheAspArgCysSerSerThrHisProhenLysTyrSerPheMetGlyAlaTyrIleCys 940
 Qy 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAGCTCCTATCCATCAAGAGACATGG 3258
 Db 941 AspaLalAysThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960
 Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGAGTTGGTTAGAGGATCTATGATGCT 3318
 Db 961 ThrThrAspaLalAysHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 Qy 3319 TCTCTAACAGTAAATATAGATATATGCGCATGGAAGATATGATATCGAGATGCTTCT 3378
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 Qy 3379 CGAGGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 23

ID AAY16738
 XX AAY16738 standard; protein; 1013 AA.

AC AAY16738;

DT 21-JUL-1999 (first entry)

XX C. trachomatis F serovar HMW protein.

KW Chlamydia; high molecular weight protein; HMW protein; urethritis;
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
 KW salpingitis; tubal occlusion; infertility; cervical cancer;
 KW arteriosclerosis; atherosclerosis.

XX Chlamydia trachomatis.

XX WO9917741-A1.

XX 15-APR-1999.

XX 01-OCT-1998; 98WO-US020737.

XX 02-OCT-1997; 97US-00942596.

XX (ANTE-) ANTEX BIOLOGICS INC.

XX Jackson JW, Pace JL;

XX WPI; 1999-287659/24.

XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
 PT cervical cancer.

XX Claim 4; Page 119-123, 141pp; English.

XX The invention relates to an isolated Chlamydia species high molecular
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
 CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
 CC be used for preventing, treating or ameliorating a disorder related to
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products
 CC can also be used for detection and diagnosis. The present sequence
 CC represents a C. trachomatis HMW protein

XX Sequence 1013 AA;

XX Alignment Scores:

Pred. No.: 0 Length: 1013
 Score: 5106.50 Matches: 980
 Percent Similarity: 98.2% Conservative: 15
 Best Local Similarity: 96.7% Mismatches: 17

Query Match: 64.8% Indels: 1
 DB: 2 Gaps: 1
 US-10-701-844-1 (1-4435) x AAY16738 (1-1013)
 Qy 382 ATGCAAGAGTCTTCCATAGTCTTCTTCAATGATCTTAGCTTATTTCTTGCTGCTCT 441
 Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
 Qy 442 TTAATGGGGGGGATATGACAGCAAAATCATGTTCTCAAGGAATTTACGATGGGAG 501
 Db 21 LeuThrGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
 Qy 502 ACGTTAACTGATCATTTCCCTATCTATAGAGATCCGAGTGGGACTACTGTTTTT 561
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 Qy 562 TCTCAGGAGAGTTAAACATTAATAAATCTTCACAAATCTTATTCAGCTTTGCTTTAAGT 621
 Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
 Qy 622 TGTTTGGGAATTTATAGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 Qy 682 GAGAACATACCGACTTCTCAAAATGGGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 Qy 742 TTTACTATTGAGGTTTAAAGATATATCCCTTTCCAAATTCATTCATTTACTTCGCTA 801
 Db 121 PheThrIleGlyGlyPheLysGlyLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
 Qy 802 CTGCTGCTGCAACGACTAATAAGGCTAGCAGACTCCGAGCAACAACATCTACACCGTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
 Qy 862 AATGTTACTATTATTTCTAAACAGACTTTTGTGTTTCTCAATTAATGAGAGTTCTCATTC 921
 Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlyLysPheSerPhe 180
 Qy 922 TATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln 200
 Qy 982 GGAATTAGCAAGCTTTGTGTTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTCT 1041
 Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys 220
 Qy 1042 CAAGTAGTCACAGTTTCTCTGCTATGCTAACGAGGCTCTTATTCCTTTGCTAGGAAT 1101
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
 Qy 1102 GTTGCAGGAGTTAAGGGGGAGGATTCCTGCTGTTCCAGGATGGGAGGAGGAGTGTCA 1161
 Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
 Qy 1162 TCATCTTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCGCGTAGAGTTT 1221
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 Qy 1222 GATGGAACTGACCGAGTAGGAGGAGGATTTACTCTTACCGGAACTGCTGTTTCTCGT 1281
 Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
 Qy 1282 AATATGGAAAACTTTGTTCTCAACATGTTGCTTCTCCTGTTTACATTCGCTGTAAG 1341
 Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
 Qy 1342 CAACCAACAGATGGAGCGCTTCTAATACGAGTAATAATTACGAGATGGAGAGCTATC 1401
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340

1402 TTCTGTAAGAATGGTGGCGCAA---GCAGGATCCCAATTAATCTCGGATCAGTTTCCTTTGAT 1458
Db
341 PheCysIysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
1459 GGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGGAAAGGGGAGCTATTAT 1518
Db
361 GlyGluGlyValAlaPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaAlaIleTyr 380
1519 GCCAAAAGCTCTCGTTTCTAATCTGCGCCTGTACAAATTTTAAGCAATATCGCTAAT 1578
Db
381 AlaIysIysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
1579 GATGTCGAGGAGATTATTTAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
Db
401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTGCCGATTTAAT 1698
Db
421 AsplleIlePheAspGlyAsnLeuIysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
1699 GCGTAACTGTCTCACAAGCCATTTCCGATCGGATCGGAGGCGAAATAACGACATTA 1758
Db
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGCGCAACGAAAT 1818
Db
461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
1819 AACGAGCAGGCGCAGCTCTTCCAAATCTTAAATAATTAACGATGGTGAAGGATACACAGGG 1878
Db
481 AsnGlnProAlaGlnSerSerGluProLeuIysIleAsnAspGlyGluGlyTyrThrGly 500
1879 GATATTGTTTGTCTAATGGAGCAGTACTTTGTACCAAAATGTTACGATAGACGACAGA 1938
Db
501 AsplleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
1939 AGGATGTTCTTCGTGAAAGGCAAAATATCAGTGAATTTCTTAAGTCAGACAGGTGG 1998
Db
521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
1999 AGTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTTCCCAACACCACCAACAG 2058
Db
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
2059 CTTCTCGGCTAACTCAGTTGATCAGCTTTTCCAACTGCAATCTGCTTCTTCTCTTTG 2118
Db
561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisSerLeuSerLeuSerLeu 580
2119 TTAGCAACAATGCAGTTACGAATCTCTACCAATCTCTCCAGCGCAAGATTTCTATCCT 2178
Db
581 LeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
2179 GCAGTCATTGGTAGCAACTCTGTTCTGTACAAATTAGTGGGCTATCTTTTGTAG 2238
Db
601 AlaValIleGlySerThrThrAlaGlyProValThrIleSerGlyProPhePheGlu 620
2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTAATCAAAAATCAAT 2298
Db
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnIysIleAsp 640
2299 GTCTCGAAATTACAGTTAGGACTAAGCCCCAGCTAATGCCCACTCAGATTGACTTA 2358
Db
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
2359 GGGAAATGAGATGCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCCTAAT 2418
Db
661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrTyrLeuAlaIleAspProAsn 680
2419 ACAGCAAAATAGTGGTCTTATCTCTGAAAGCTACATGAGCTTAAACTGGGTATATCTCT 2478
Db
681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700
2479 GGGCCTGACGAGTAGTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538

701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720
2539 CGATCTCGCAATTCAGCAATTCAGCAAGTGTGTGATGGCGCTCTATTATTGTCGAGATT 2598
Db
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
2599 TGGGTTTCTCGAGATTTTCGAATTTCTTATCATGACCCCGATGCTTTTAGTCAGGATAT 2658
Db
741 TrpValSerGlyValSerAsnPheSerTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
2659 CGGTATATTAGTGGGGTTTATTCCTTAGGACCAACTCTACTTTGGATCATCGATGTT 2718
Db
761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTGTTCCCAT 2778
Db
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCCCAACAGCTTTATGTGATCCTAT 2838
Db
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
2839 TTGTTTCGAGATGCTTTTATCCGTCCTAGCTACCGGTTTGGAAATCAGCATATGAACCC 2898
Db
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
2899 TCATATACATTTTCGAGGAGGAGCGATCTTCGTCGGGATAATACTGTCTGGCTCGAGAG 2958
Db
841 SerTyrThrPheAlaGluGluSerAspValArgThrPaspAsnAsnCysLeuValGlyGlu 860
2959 ATTGAGCGGGATTTACCGATTTGTGATTACTCCATCTAAGCTCTATTGAATGAGTTGCT 3018
Db
861 IleGlyValGlyLeuProIleValThrThrProSerLysLeuTyrLeuAsnGluLeuArg 880
3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGAGCGAT 3078
Db
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
3079 CAAGCTCGGCATTTCAAGAGCGGACATCTCCTAAATCTATCAGTTCTCTGTCGAGTGAAG 3138
Db
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
3139 TTGTGATCATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGGCGCTTATATCTGT 3198
Db
921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAACGCTCTCTATCCCATCAAGACATGG 3258
Db
941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGTTGTGGTTAGAGGATCTATGTATGCT 3318
Db
961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
3319 TCTCTAAACAAGTAATATAGAATATATGGCCATATGGAAGATATGAGTATCAGATGCTTCT 3378
Db
981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
3379 CGAGCTATGGTTTGAGTCAGGAAGTAGAGTCCGGTTC 3417
Db
1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 24
AAG83207
ID AAG83207 standard; protein; 1006 AA.
XX
AC AAG83207;
XX 05-SEP-2001 (first entry)
XX DE Protein encoded by Chlamydia trachomatis pmpG gene.
XX

KW Chlamydia; vaccine; infection; fusion protein; antigen; 891
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease; 164
 KW acute respiratory tract infection; Cap; CT529; OMCB; 951
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA. 184
 XX Chlamydia trachomatis. 1011
 XX W0200140474-A2. 204
 XX 07-JUN-2001. 1071
 XX 04-DEC-2000; 2000WO-US032919. 224
 XX 03-DEC-1999; 99US-00454684. 1131
 XX 19-APR-2000; 2000US-00556877. 244
 XX 20-JUN-2000; 2000US-00598419. 1191
 XX (CORI-) CORIXA CORP. 264
 XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J; 1251
 XX MPI; 2001-374831/39. 284
 XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic 1311
 XX inflammatory disease, trachoma, acute respiratory tract infections, 304
 XX atherosclerosis and heart disease. 1371
 XX Claim 2; Page 214-216; 295pp; English. 324
 XX The present sequence is provided in a specification relating to compounds 1431
 XX and methods for the treatment and diagnosis of chlamydial infection. The 344
 XX compounds provided include polypeptides and fusion proteins comprising 1491
 XX immunogenic portions of Chlamydia antigens and DNA sequences encoding 364
 XX such polypeptides. They are useful for vaccinating against chlamydial 1551
 XX infection, which causes pelvic inflammatory disease, trachoma, acute 384
 XX respiratory tract infections, atherosclerosis and heart disease 1611
 XX Sequence 1006 AA; 404
 XX Alignment Scores: 1671
 XX Pred. No.: 0 Length: 1006
 XX Score: 5090.00 Matches: 979
 XX Percent Similarity: 99.9% Conservative: 2
 XX Best Local Similarity: 99.7% Mismatches: 1
 XX Query Match: 64.6% Indels: 0
 XX DB: 4 Gaps: 0
 XX US-10-701-844-1 (1-4435) x AAG83207 (1-1006)
 QY 472 ATGGTTCCTCAAGGAATTTACGATGGGAGAGGTTAACTGTATCATTTCCCTATACTGTT 531
 Db 25 MetileProGInGlyIleYrAspGlyGluThrLeuThrValSerPheProfyrThrVal 44
 QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCTCGAGGAGAGTTAAACATTTAAAAATCTT 591
 Db 45 IieGlyAppProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64
 QY 592 GACAAATCTTATTCGAGCTTTCCTTTAAAGTTGCTTTTGGGAATTTATAGGGAGTTTACT 651
 Db 65 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84
 QY 652 GTTTTAGGAGGAGGACCTGTTGACTTTCGAGACATACGACTTCTACAAATGGGCA 711
 Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104
 QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTGAGGGTTTTAAAGAATTTATCC 771
 Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124
 QY 772 TTTTCCAATTCGAATTCATTACTGCGGTATGCGCTGCTGCAACGACTAATAAGGTAC 831
 Db 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144

832 CAGACTCCGACGACAAATCTACCGTCTAATGGTACTATTTATTTCTAAACAGATCTT 891
 Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleYrSerLysThrAspLeu 164
 QY 892 TTGTTACTCAATAATGAGAAGTTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGGA 951
 Db 165 LeuLeuLeuAsnAsnGluLysPheSerPheTySerAsnLeuValSerGlyAspGlyGly 184
 QY 952 GCTATAGATGCTAAGAGCTTAAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAAAGAA 1011
 Db 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204
 QY 1012 AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCT 1071
 Db 205 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224
 QY 1072 AACGAGCTCTTATTTGCTTTGTAGGAATTTGTCAGGAGTAAAGAGGGGAGGATGCT 1131
 Db 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244
 QY 1132 GCTGTTCAAGATGGGACGAGGAGTGTCTCATCTTCACTTCAACAGAAATCCAGTAGTA 1191
 Db 245 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal 264
 QY 1192 AGTTTTTCCAGAAATACTGCGGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGGG 1251
 Db 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 284
 QY 1252 ATTTACTCTACGGGAACGTTGCTTCTGTAATATGGAATAACCTGTTTCTCAACAAT 1311
 Db 285 IleYrSerYrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 304
 QY 1312 GTTCCTCTCTGTTTACATTTGCTTAAGCAACCAACAGTGACAGGCTTCTAATACG 1371
 Db 305 ValAlaSerProValYrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324
 QY 1372 AGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAAGATGTGTGCGCAACGAGGATCC 1431
 Db 325 SerAsnAsnYrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344
 QY 1432 AATAACTCTGATCAGTTTCTGTTGATGGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491
 Db 345 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 364
 QY 1492 GCTGCTGGGAAAGGGAGCTATTTATGCCAAAAGCTCTCGGTTGTAACTGTGCGCCT 1551
 Db 365 AlaAlaGlyLysGlyGlyAlaIleYrAlaLysLysLeuSerValAlaAsnCysGlyPro 384
 QY 1552 GTACAAATTTTAAAGGAATATCGCTAATCATGGTGGAGCGATTTATTTAGGAGAATCTGA 1611
 Db 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleYrLeuGlyGlySerGly 404
 QY 1612 GAGCTCAGTTTATCTGCTGATTTAGGATATTTTTCGATGGGAATCTTAAAGAAACA 1671
 Db 405 GluLeuSerLeuSerAlaAspYrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424
 QY 1672 GCCAAGAGAAATGCTGCCGATGTTAATGCGGTAACTGTGCTCTCAAGACCAATTCGATG 1731
 Db 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444
 QY 1732 GGATCGGAGGGAATAATACGACATTAAGAGCTAAAGCGGCGCATCAGATCTCTTTAAT 1791
 Db 445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464
 QY 1792 GATCCCATCGAGATGGCAACCGAATAACACGACGCGCAGCTTCTCCAACTTCTAAAAA 1851
 Db 465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484
 QY 1852 ATTTAACGATGTGAAGGATACACAGGGGATATTTGTTTTGCTTAATGAAGACAGTACTT 1911
 Db 485 IleAsnAspGlyGlyGlyYrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504
 QY 1912 TACCAAAATGTTTACGATAGACAGGAGGATTTGTTCTTCGTGAAAAGGCAAAATTTATCA 1971

Db 505 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLeuYsAlaYsLeuSer 524
 QY 1972 GTGAATTCCTAAGTCAGACAGGTGGAGTCTGTATATAGGAAGCTGGGAGTACATGGGAT 2031
 Db 525 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 544
 QY 2032 TTGTAACTCCACACACCACACACAGCTCTGCGCTAAATCAGTGTATGATCAGCTTCC 2091
 Db 545 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 564
 QY 2092 AATCTGCATTGTCTCTTCTCTTCTTGTAGCAAAACAATGCAATGCAATCTCTCTACC 2151
 Db 565 AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 584
 QY 2152 AATCTCCAGCGCAAGATTCATCTCTGAGTCAATGGTAGCAAACTCTGCTGTCTGT 2211
 Db 585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 604
 QY 2212 ACAATTAGTGGGCTATCTTTTGGAGATTTGGATGATACAGCTTATGATAGGTATGAT 2271
 Db 605 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 624
 QY 2272 TGGCTAGGTTCTAATCAAAAAATCAATGCTCTGAAATTAAGTTAGGGACTAAGCCCCCA 2331
 Db 625 TrpLeuGlySerAsnGlnYsIleAsnValLeuYsLeuGlnLeuGlyThrYsProPro 644
 QY 2332 GCTAATGCCCCATCAGATTGACTTAGGGAATGAGATGCTAAGTAGTGTATCAAGGA 2391
 Db 645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProYsYrGlyThrGlnGly 664
 QY 2392 AGCTGAAGCTTCGCTGGGATCTTAATACAGCAAAATTAAGTCTTATCTCTGAAGCT 2451
 Db 665 SerTrpYsLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProYrThrLeuYsAla 684
 QY 2452 ACATGACTAAACTGGGTATATCTGGGCTGAGCGAGTAGCTTCTTTGGTTCCAAAT 2511
 Db 685 ThrTrpYsThrGlyYrAsnProGlyProGluArgValAlaSerLeuValProAsn 704
 QY 2512 AGTTATGGGGATCCATTTTATAGATACGATCTGCGCATTCAGCAATTCAGCAAGTGTG 2571
 Db 705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
 QY 2572 GATGGGCTCTTATGTCAGGATTAATGGTTCTGGAGTTTCGAAATTCCTCTATCAT 2631
 Db 725 AspGlyArgSerYrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheYrHis 744
 QY 2632 GACCGGATCTTTAGTTCAGGATATCGGTATATTAGTGGGTTATCTCTTAGGAGCA 2691
 Db 745 AspArgAspAlaLeuGlyGlnGlyYrArgYrIleSerGlyGlyYrSerLeuGlyAla 764
 QY 2692 AACTCTTACTTTGGATCATCGATTTGGTCTAGCATTTACCGAAGTATTGGTAGATCT 2751
 Db 765 AsnSerYrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
 QY 2752 AAGATTAATGATGTGTGTTCCATCATCATCTGTCATAGATCCGTTTATCTATCT 2811
 Db 785 LysAspYrValValCysArgSerAsnHisIleAlaCysIleGlySerValYrLeuSer 804
 QY 2812 ACCCAAGCTTTATGTGATCTTATTTGTCGGAGATCGTTTATCTCGGTAGCTAC 2871
 Db 805 ThrGlnGlnAlaLeuCysGlySerYrLeuPheGlyAspAlaPheIleArgAlaSerYr 824
 QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAGCGATGTTCTGT 2931
 Db 825 GlyPheGlyAsnGlnHisMetYsThrSerYrThrPheAlaGluSerAspValArg 844
 QY 2932 TGGGATTAATCTCTGCTGGAGAGATTGGAGCGGGATTAACGATGTGATTACTCCA 2991
 Db 845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864
 QY 2992 TCTAAGCTCTATTGATGATGCTGCTCTTCTGCAAGCTGAGTTTCTTATGCGCAT 3051

865 SerYsLeuYrLeuAsnGluLeuArgPropheValGlnAlaGluPheSerYrAlaAsp 884
 3052 CATGAATCTTTTACAGAGAGAGGCGCATCAAGCTCGGCGATTCAAGAGCGGAGCATCTCCTA 3111
 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheYsSerGlyHisLeuLeu 904
 3112 AATCTATCAGTTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAATAA 3171
 905 AsnLeuSerValProValGlyValYsPheAspArgCysSerSerThrHisProAsnYs 924
 3172 TATAGCTTTATGCGCGCTTATATCTGATGCTTATCGCACCCTCTCTGTTACTGAGACA 3231
 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
 3232 ACGCTCTTATCCCATCAAGAGACATGGCAACAGATGCTCTTTCATTTCATTAGCAAGATGGA 3291
 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964
 3292 GTTGTGGTTAGAGGATCTATGATGCTTCTCTTAAACAAGTAATATAGAAATATATGGCCAT 3351
 965 ValValValArgGlySerMetYrAlaSerLeuThrSerAsnIleGluValYrGlyHis 984
 3352 GGAAGATATGATATCGAGATGCTCTCGAGGCTATGTTTTCAGTGCAGAGATGAGTTC 3411
 985 GlyArgYrGluTyrArgAspAlaSerArgGlyYrGlyLeuSerAlaGlySerYsVal 1004
 3412 CGGTTC 3417
 1005 ArgPhe 1006
 RESULT 25
 ID ABB94178 standard; protein; 1006 AA.
 AC ABB94178;
 XX 05-JUN-2002 (first entry)
 XX Chlamydia protein sequence SEQ ID NO:190.
 DE Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.
 XX Chlamydia sp.
 OS WO200208267-A2.
 PN 31-JAN-2002.
 XX 20-JUL-2001; 2001WO-US023121.
 PF 20-JUL-2000; 2000US-00620412.
 PR 23-APR-2001; 2001US-00841132.
 XX (CORI-) CORIXA CORP.
 PA Fling SP, Skeiky YAW, Probst P, Bhatia A;
 PI WPI; 2002-179901/23.
 DR Novel compositions comprising Chlamydia Cap1 protein and its use in the
 PT treatment of Chlamydia infection.
 XX Example 1; Page 245-247; 537pp; English.
 PS The present invention describes compositions comprising a Chlamydia Cap1
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.
 CC Chlamydia DNA and protein sequences from the present invention can have
 CC antibacterial and immunostimulant activities, and can be used in
 CC vaccines. Compounds from the present invention can be used for eliciting
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell
 CC response or inhibiting the development of a Chlamydia infection in an

CC animal. Methods from the present invention can be used: for detecting the
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia
 CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 CC sequences used in the exemplification of the present invention
 XX

SQ Sequence 1006 AA;

Alignment Scores:
 Pred. No.: 0 Length: 1006
 Score: 5090.00 Matches: 979
 Percent Similarity: 99.9% Conservative: 2
 Best Local Similarity: 99.7% Mismatches: 1
 Query Match: 64.6% Indels: 0
 DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) x ABB94178 (1-1006)

QY	472	ATGTTCTCTCAAGGAATTACGATGGGAGAGCTTAACGTGATCATTTCCCTACTGTT	531
DB	25	MetIleProGInGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal	44
QY	532	ATAGAGAGATCCGAGTGGGACTACTGTTTTCTCGCAGGAGCTTAACATTAAAAATCTT	591
DB	45	IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLeuSerLeu	64
QY	592	GACAAATCTATTGCGAGCTTGCTTTTAAGTTGTTTTGGGAACTTATTAGGGAGTTTTACT	651
DB	65	AspAenSerIleAlaAlaLeuProLeuSerCysPheGlyAenLeuLeuGlySerPheThr	84
QY	652	GTTTTAGGGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTACAAATGGGCA	711
DB	85	ValLeuGlyArgGlyHisSerLeuThrPheGluAenIleArgThrSerThrAenGlyAla	104
QY	712	GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGGTTTTAAAGAATTATCC	771
DB	105	AlaLeuSerAenSerAlaAlaAspGlyLeuPheThrIleGluGlyPheGlySerLeuSer	124
QY	772	TTTTTCCAATTTGCAATTTCAATTTCTGCGTACTGCTGCTGCAACGACTTAATAAGGGTAGC	831
DB	125	PheSerAenCysAenSerLeuLeuAlaValLeuProAlaAlaThrThrAenLysGlySer	144
QY	832	CAGACTCGACACACATCTACACGCTCTAATGTTACTATTATTCTTAAACAGACTCTT	891
DB	145	GlnThrProThrThrSerThrProSerAenGlyThrIleTyrSerLysThrAspLeu	164
QY	892	TTGTACTCAATAAGAGAGTTCTCATTTAGTAAATTTAGTCTCTGGAGATGGGA	951
DB	165	LeuLeuLeuAenAenGlnLysPheSerPheTyrSerAenLeuValSerGlyAspGlyGly	184
QY	952	GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTTAGCAAGCTTTGTCTTCCAAGAA	1011
DB	185	AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu	204
QY	1012	AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCT	1071
DB	205	AenThrAlaGlnAlaAspGlyAlaCysGlnValThrSerPheSerAlaMetAla	224
QY	1072	RACGAGGCTCTATTGCTTTCTAGCGAATGTTGACAGGATTAAGAGGGGAGGATGCT	1131
DB	225	AenGluAlaProIleAlaPheValAlaAenValAlaGlyValArgGlyGlyIleAla	244
QY	1132	GCTGTTTCAGGATGGGAGGAGTGTCTATCATCTTCTTCAACAGAAGATCCAGTAGTA	1191
DB	245	AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrThrGluAspProValVal	264
QY	1192	AGTTTTCCAGAAATCTACGGTAGGTTGATGGGAAGCTAGCCCGAGTAGGAGAGGG	1251
DB	265	SerPheSerArgAenThrAlaValGluPheAspGlyAenValAlaArgValGlyGly	284
QY	1252	ATTTACTCTACGGAGCTGCTTTCTTCGTAATTAATGGAAGAACCTTTCTCAACAT	1311
DB	285	IleTyrSerTyrGlyAenValAlaPheLeuAenAenGlyLysThrLeuPheLeuAenAen	304

QY	1312	GTTCGTTCTCTGTTTACATTTGCTGTAAGCAACCAAGTGGACAGGCTTCTTAATACG	1371
DB	305	ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAenThr	324
QY	1372	AGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAAGATGTTGGCAGCAGGATCC	1431
DB	325	SerAenAenTyrGlyAspGlyGlyAlaIlePheCysLysAenGlyAlaGlnAlaGlySer	344
QY	1432	AATAACTCTGATCAGTTTCTTTGATGGAGGAGTAGTTTCTTTTAGTAGCAATGTA	1491
DB	345	AenAenSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAenVal	364
QY	1492	GCTGCTGGAAAGGGGAGCTATTATTATGCCAAAAAGCTCTCGGTGCTTAACCTGGCCCT	1551
DB	365	AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAenCysGlyPro	384
QY	1552	GTACAAATTTTAAAGGAATATCGCTAATGATGGTGGAGGATTTATTTAGGAGAATCTGA	1611
DB	385	ValGlnPheLeuAenAenIleAlaAenAspGlyGlyAlaIleTyrLeuGlyGluSerGly	404
QY	1612	GAGCTCAGTTTATCTGCTGATTTATGGAGATATTATTTTCGATGGGAATCTTTAAAGAACA	1671
DB	405	GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAenLeuLysArgThr	424
QY	1672	GCCAAAGAGAATGCTGCGATGTTAATGGCGTAATCTGTCTCTCAAGCCATTTTCGATG	1731
DB	425	AlaLysGluAenAlaAlaAspValAenGlyValThrValSerSerGlnAlaIleSerMet	444
QY	1732	GGATCGGAGGGAATAACGACATTAAGACTTAAGCAGGAGGATTAAGGAGGATCTCTTAAT	1791
DB	445	GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAen	464
QY	1792	GATCCATCGAGATGGCAACGAAATACCGAGCAGCGAGCTTCTTCAAACTTCTTAAAA	1851
DB	465	AspProIleGluMetAlaAenGlyAenAenGlnProAlaGlnSerSerLysLeuLys	484
QY	1852	ATTACAGATGTTGAAGATACACAGGGGATATTGTTTTGCTTAATGGAAGCAGTCTTTG	1911
DB	485	IleAenAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAenGlySerSerThrLeu	504
QY	1912	TACCAAAATGTTACGATAGACAGGAGGATTTCTTCTGTAAGGCAAAATATCA	1971
DB	505	TyrGlnAenValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer	524
QY	1972	GTGAATCTCTAAGTCAGACAGGAGTCTGTATATGGAAGCTGGGAGTACATGGAT	2031
DB	525	ValAenSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp	544
QY	2032	TTTGTAACTCCAAACCAACCAAGCCTCTGCGCTCTGCTTAATTCAGTTGATCAACGCTTCC	2091
DB	545	PheValThrProGlnProGlnProGlnProAlaAlaAenGlnLeuIleThrLeuSer	564
QY	2092	AATCTGCAATTTGCTCTTCTTCTTGTAGCAACAAATGAGTGTACCAATCTCTCTACC	2151
DB	565	AenLeuHisLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThrAenProThr	584
QY	2152	AATCTCCAGCGCAAGATTTCTCATCTGCACTTGTGTAGCACAACCTGCTGTTTCTGTT	2211
DB	585	AenProProAlaGlnAenSerHisProAlaValIleGlySerThrThrAlaGlySerVal	604
QY	2212	ACAATTAGTGGGCTTATCTTTTGTAGGATTTGGATGATACAGCTTATGATAGTATGAT	2271
DB	605	ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp	624
QY	2272	TGGCTAGCTTCTTAATCAAAAATCAATCTCTGGAATTTACAGTGTAGGACCTAAGCCCCA	2331
DB	625	TrpLeuGlySerAenGlnLysIleAenValLeuLysLeuGlnLeuGlyThrLysProPro	644
QY	2332	GCTAATGCCCCATCAGATTTGACTTAGGGAATGAGATGCTTAACTATGGCTTATCAAGA	2391
DB	645	AlaAenAlaProSerAspLeuThrLeuGlyAenGluMetProLysTyrGlyTyrGlnGly	664

QY 2392 AGCTGGAGCTTGGCTGGATCCTTAATACAGCAAAATAATGGTCTTATCTCTGAAGCT 2451
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684
QY 2452 ACATGGACTAAAGCTGGGTATATCTCGGCCCTGAGCGAGTAGCTTCTTTGGTCCAAAT 2511
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704
QY 2512 AGTTATGGGGATCCATTTTATAGATATACGATCGGCATTCAGCAATTCAGCAAGTGG 2571
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
705 SerLeuTrpGlySerLeuAlaAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
QY 2572 GATGGCGCTCTATTATGTCGAGGATATGGGTTCTCGAGTTTCGAGTTTCTTCTATCAT 2631
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 744
QY 2632 GACCGCGATGCTTTAGTCAGGATATCGGTATATTAGTGGGGTTATTCTTAGAGCA 2691
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
QY 2692 AACTCTACTTTGGATCATCGATGTTGGTCTAGCATTTTACCGAGTATTGGTAGATCT 2751
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
765 AsnSerTyrPheGlySerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
QY 2752 AAGATTATGATGTGCTGCTCCAAATCATCATGCTTCATAGATCCGTTTATCTATCT 2811
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
785 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 804
QY 2812 ACCCAACAAAGCTTTATGTCGATCCTATTGTTGCGAGATCGGTTTATCCGTGCTACTAC 2871
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824
QY 2872 GGGTTGGGAATCAGCATATGAAACCTATATACATTTCCAGAGAGCGGATGTCGT 2931
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluSerAspValArg 844
QY 2932 TCGGATAATTAACGTCTGGCTGAGAGATGGAGCGGATTACCGATTGATGTTACTCCA 2991
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864
QY 2992 TCTAAGCTCTATTGATGATGCTGCTCTCTTCGTCAGAGCTGAGTTTCTTATCCGAT 3051
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884
QY 3052 CATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTA 3111
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
QY 3112 AATCTATCAGTTCTGTTGGAGTCAAGTTTGTATCGATGTTTCTAGTACACATCCTAATAA 3171
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisLysProAsnLys 924
QY 3172 TATAGCTTTATGGCGCTTATCTGTGATGCTTATCGCCACCTCTCTGTTACTAGACA 3231
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
QY 3232 ACGTCTCTATCCCATCAAGAGACATGACACAGATCCCTTCTATTAGCAAGACATGGA 3291
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964
QY 3292 GTTGTGTTAGAGATCTATGATGCTTCTCAACAGTAAATATAGAGTATATGGCCAT 3351
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984
QY 3352 GGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGTAGTCAGGAGTAGATC 3411
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004
QY 3412 CGGTTT 3417
DB |||||||
1005 ArgPhe 1006
DB |||||||

RESULT 26

AAB13633
ID AAB13633 standard; protein; 982 AA.
AC AAB13633;
XX
XX 02-FEB-2001 (first entry)
XX
XX C. trachomatis pmpG gene protein.
XX
XX Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease; antibacterial.
XX
XX Chlamydia trachomatis.
XX
XX Key Location/Qualifiers
FT Misc-difference 981
FT /note= "Unspecified amino acid"
XX
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-US029012.
XX
XX 08-DEC-1998; 98US-00208277.
PR 08-APR-1999; 99US-00288594.
PR 01-OCT-1999; 99US-00410568.
PR 22-OCT-1999; 99US-00426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
PI WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
PT acid sequence encoded by polynucleotide sequence.
XX
XX Claim 2; Page 181-184; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamidia are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention
XX
XX SQ Sequence 982 AA;
XX
XX Alignment Scores:
Pred. No.: 0 Length: 982
Score: 5084.00 Matches: 978
Percent Similarity: 99.8% Conservatives: 2
Best Local Similarity: 99.6% Mismatches: 2
Query Match: 64.5% Indels: 0
DB: 3 Gaps: 0
US-10-701-844-1 (1-4435) x AAB13633 (1-982)
QY 472 ATGGTTCCTCAAGAAATTTACGATGGGAGCGTTAACTGATATCTTCCTACTACTGTT 531
DB |||:|||||
1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20


```

Db 741 AnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
QY 2752 AAAGATTATGATGTCGTTCCCAATCATCATGCTTGCATAGGATCGTTTATCTATCT 2811
Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
QY 2812 ACCCAACAGCTTATGTCGATCCTATTTGTCGGAGATGCGTTTATCGTGCTAGCTAC 2871
Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800
QY 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGCATGTCGT 2931
Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820
QY 2932 TGGGATAATAACTGCTGCTGGAGATGGAGCGGATTTACCGATTTGTGATTAATCTCCA 2991
Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
QY 2992 TCTAAGCTCTATTTGAATGATGTCGCTCTCTTTCGTCGACGTCAGTCTTCTTATGCGAT 3051
Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
QY 3052 CATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCGATTTCAAGAGCGGACATCTCCTA 3111
Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
QY 3112 AATCTATCACTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTATAATA 3171
Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
QY 3172 TATAGCTTATGCGGCTTATCTGTCGATGCTTATCGACCATCTCTGCTACTGAGACA 3231
Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920
QY 3232 ACGCTCTATCCATCAAGAGACATGACCAACAGATGCCCTTTCATTTAGCAAGACATGGA 3291
Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAlaGlyHisGly 940
QY 3292 GTTGTGGTTAGAGATCTATGATGCTTCTTCAAGATATATAGATATATGCGCAT 3351
Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960
QY 3352 GGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTGAGTGCAGGAGCTAGAGTC 3411
Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980
QY 3412 CGGTTTC 3417
Db 981 ***Phe 982

```

RESULT 27

AAG83201

ID AAG83201 standard; protein; 982 AA.

XX AC AAG83201;

DT 05-SEP-2001 (first entry)

XX DE Protein encoded by Chlamydia trachomatis pmpG gene.

XX KW Chlamydia; vaccine; infection; fusion protein; antigen;

KW KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;

KW KW acute respiratory tract infection; Cap1; CT529; OMCB;

KW KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.

OS Chlamydia trachomatis.

XX WO200140474-A2.

XX PN 07-JUN-2001.

XX PD 04-DEC-2000; 2000WO-US032919.

XX PF

```

XX 03-DEC-1999; 99US-00454684.
PR 19-APR-2000; 2000US-00556877.
PR 20-JUN-2000; 2000US-00598419.
XX (CORI-) CORIXA CORP.
XX Probat P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX WPI; 2001-374831/39.
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease.
XX Claim 2; Page 189-191; 295pp; English.
XX The present sequence is provided in a specification relating to compounds
XX and methods for the treatment and diagnosis of chlamydial infection. The
XX compounds provided include polypeptides and fusion proteins comprising
XX immunogenic portions of Chlamydia antigens and DNA sequences encoding
XX such polypeptides. They are useful for vaccinating against chlamydial
XX infection, which causes pelvic inflammatory disease, trachoma, acute
XX respiratory tract infections, atherosclerosis and heart disease
XX SQ Sequence 982 AA;

```

Alignment Scores:

```

Pred. No.: 0 Length: 982
Score: 5084.00 Matches: 978
Percent Similarity: 99.8% Conservative: 2
Best Local Similarity: 99.6% Mismatches: 2
Query Match: 64.5% Indels: 0
DB: 4 Gaps: 0

```

US-10-701-844-1 (1-4435) x AAG83201 (1-982)

```

QY 472 ATGTTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTACTACTGT 531
Db 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20
QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTTCTGCGAGGAGAGTTAACTAAAAATCTT 591
Db 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40
QY 592 GACAACTTCTATTCACGCTTTCCTTTAACTGTTGTTGGGAACCTATTAGGAGTCTTACT 651
Db 41 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60
QY 652 GTTTTAGGGAGGAGACACTGCTTGTGACTTTTCGAGAACATACGAGCTTCTACAAATGGGCA 711
Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
QY 712 GCTCTAAGTAATAGCGCTGCTGATGGAGCTGTTTACTATTATGAGGTTTAAAGAAATATCC 771
Db 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 100
QY 772 TTTTCAATTGCAATTCATTACTTTCGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
Db 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120
QY 832 CAGACTCCGACGACACATCTACACGCTCTAAATGGTACTATTTATTTCTAAACAGATCTT 891
Db 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleThrSerLysThrAspLeu 140
QY 892 TTGTTACTCAATAATAGAAAGTTCTCATCTATATAGTAATTTAGTCTCTGAGATGGGGA 951
Db 141 LeuLeuLeuAsnAsnGluLysPheSerPheThrSerAsnLeuValSerGlyAspGlyGly 160
QY 952 GCTATAGATGCTAAGACTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCCAGAA 1011
Db 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180

```

QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCTAGTAGTACCAAGTTTCTCTGCTATGGCT 1071
DB 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSerAlaMetAla 200
QY 1072 AACGAGGCTCCTATTGCTTTGTACGGAATGTTGCAGGAGTAAGAGGGGAGGAGTTCCT 1131
DB 201 AsnGlnAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 220
QY 1132 GCTGTTACGATGGCAGCAGGAGTGTCTCATCTACTTCAACAGAAGATCCAGTAGTA 1191
DB 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal 240
QY 1192 AGTTTTCAGAAATACCTGCGGTAGACTTTGATGGGAACGTAGCCGAGTAGGAGGAGG 1251
DB 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 260
QY 1252 ATTTACTCTACGGGAAGTTCCTTCTGTAATAATGGAATAACCTTGTTCTCAACAAT 1311
DB 261 IleThrSerThrGlyAsnValAlaPheLeuAsnGlyValThrLeuPheLeuAsn 280
QY 1312 GTTGCTTCTCTGTTTACATTCCTGCTAAGCAACCAACAGTGGACAGCTTCTTAATACG 1371
DB 281 ValAlaSerProValTyrrileAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300
QY 1372 AGTAATAATTACGGAGATGGAGAGCTATCTTCTGTAAAGATGGTGGCAAGCAGATCC 1431
DB 301 SerAsnAsnThrGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320
QY 1432 AATAACTCTGATCAGTTCTCTTTCATGGAGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491
DB 321 AsnAsnSerGlySerValSerPheAspGlyGluGlyValAlaPhePheSerSerAsnVal 340
QY 1492 GCTGCTGGAAAGGGGAGCTATTATGCCAAAAGCTCTCGTTGCTAACTGTGGCCCT 1551
DB 341 AlaAlaGlyLysGlyAlaIleTyrrAlaLysLysLeuSerValAlaAsnCysGlyPro 360
QY 1552 GTACAAATTTTAAAGAAATACCTTAATGATGGTGGAGCGATTATTTAGGAGAATCTGA 1611
DB 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyAlaIleTyrrLeuGlyGluSerGly 380
QY 1612 GAGCTCAGTTTCTGCTGATTCGAGATATTATTTTCGATGGGAATCTTAAAGAACAA 1671
DB 381 GluLeuSerLeuSerAlaAspTyrrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400
QY 1672 GCCAAAGAGATGCTGCCGATGTTAATGGCGTAACCTGTCTCACAGCCATTTGATG 1731
DB 401 AlaLysGluAsnAlaAlaAspValAlaAsnGlyValThrValSerSerGlnAlaIleSerMet 420
QY 1732 GGATCGGAGGGAATAATAACGACATTAGAGCTAAAGCGAGCGATCAGATTCTCTTAAT 1791
DB 421 GlySerGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440
QY 1792 GATCCCATCGAGTGGCAACGGAATAACACGAGCGCAGTCTTCCAAACTCTTAAAA 1851
DB 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys 460
QY 1852 ATTAACGATGGTGAAGGATACACAGGGGATTTGTTTTGCTTAATGGAAGCAGTACTTTG 1911
DB 461 IleAsnAspGlyGluGlyTyrrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
QY 1912 TACCAAAATGTTACGATAGACGAGGAGGATTTGTTCTTCGTGAAAGGCAAAATATCA 1971
DB 481 TyrrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500
QY 1972 GTGAATCTCTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGGAT 2031
DB 501 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrrMetGluAlaGlySerThrLeuAsp 520
QY 2032 TTTGTAACTCCAAACCAACCAACAGCTCTCGCGCTTAATCAGTTGATCAGCTTTCC 2091
DB 521 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 540
QY 2092 AATCTGCATTTGCTCTTCTTCTTGTGTTAGCAACAAATGCGAGTTACGAATCTCTCTACC 2151

DB 541 AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560
QY 2152 AATCTCCAGCGCAGAGATTTCTCATCTCTGCACTTATGGTAGCACAACTGCTGTTCTGTT 2211
DB 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580
QY 2212 ACAATTAGTGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGTATGAT 2271
DB 581 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrrAspArgTyrrAsp 600
QY 2272 TGGCTAGTCTTAATCAAAAATCAATGCTCTGAAATTTACAGTTTAGGCACTAAAGCCCCA 2331
DB 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
QY 2332 GCTAATGCCCATCAGATTTTGACTCTAGGGGAATGAGATGCTTAAGTATGGCTATCAAGA 2391
DB 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrrGlyTyrrGlnGly 640
QY 2392 AGCTGGAAAGCTTGGTGGGATCTTAATACAGCAAAATAATGCTCTTATCTACTCTGAAGCT 2451
DB 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrrThrLeuLysAla 660
QY 2452 ACATGGGACTAAACTGGCTATAATCTCGGCTGAGCGAGTAGCTTCTTTGGTTCCAAAT 2511
DB 661 ThrTrpThrLysThrGlyTyrrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
QY 2512 AGTTTATGGGATCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTG 2571
DB 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
QY 2572 GATGGGCTCTTATTTGTGAGGATTTATGGGTTCTGGAGTTTCAATTTCTTCTATCAT 2631
DB 701 AspGlyArgSerTyrrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrrHis 720
QY 2632 GACCGGATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTATCTCTTAGGACA 2691
DB 721 AspArgAspAlaLeuGlyGlnGlyTyrrArgTyrrIleSerGlyGlyTyrrSerLeuGlyAla 740
QY 2692 AACTCTACTTTGATCATCATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751
DB 741 AsnSerTyrrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
QY 2752 AAAGATTATGATGTTGTCGTTCCATCATCATGCTTCCATAGGATCCGTTTATCTATCT 2811
DB 761 LysAspTyrrValValCysArgSerAsnHisAlaCysIleGlySerValTyrrLeuSer 780
QY 2812 ACCCAACAGCTTTATGAGTTCCTATTGTTGCGAGATGCGTTTATCCGTCTAGCTAC 2871
DB 781 ThrGlnGlnAlaLeuCysGlySerTyrrLeuPheGlyAspAlaPheIleArgAlaSerTyrr 800
QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGAGAGCGATGTTCT 2931
DB 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrrThrPheAlaGluGluSerAspValArg 820
QY 2932 TGGGATAATAACTGTTCTGCTGGAGAGATTGGAGCGGATTTACCGATTGTTGATTACTCCA 2991
DB 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
QY 2992 TCTAAGCTCTATTGCAATGAGTTGCTCTTCTGTCGCAAGCTGAGTTTCTTATCCGAT 3051
DB 841 SerLysLeuTyrrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrrAlaAsp 860
QY 3052 CATGAATCTTTTACAGAGAAAGCGATCAAGCTCGGGCATTCAGAGCGGACATCTCTTA 3111
DB 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
QY 3112 AATCTATCAGTTCTGTTGGAGTGAAGTTTGATTCGATGTTCTAGTACACATCTTAATAAA 3171
DB 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
QY 3172 TATAGCTTTATGGCGGCTTATATCTGATGCTTATCGCACCATCTCTGTTGATCTAGACA 3231

Db 901 TyrSerPheMetAlaAlaTyrlleCysAspAlaTyArgThrIleSerGlyThrGluThr 920
 QY 3232 ACGTCTCTATCCATCAAGACATGACGACAAAGATGCTTTCATTTACCAAGACATGGA 3291
 Db 921 ThrLeuLeuSerHieGlnGluThrTrpThrThrAspAlaPheHieLeuAlaArgHieGly 940
 QY 3292 GTTGTGGTTAGAGGATCTATGATGCTTCTCTAAACAAGTAATATAGACGATATATGCCAT 3351
 Db 941 ValValValArgGlySerMetTyAlaSerLeuThrSerAsnIleGluValTyGlyHie 960
 QY 3352 GGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTTCAGTCGAGGAAGTAGAGTC 3411
 Db 961 GlyArgTyArgAlaSerArgGlyTyArgGlyLeuSerAlaGlySerIysVal 980
 QY 3412 CGGTTT 3417
 Db 981 ***Phe 982

RESULT 28
 ID ABB94172 standard; protein; 982 AA.
 AC ABB94172;
 DT 05-JUN-2002 (first entry)
 XX Chlamydia protein sequence SEQ ID NO:176.
 DE Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.
 XX Chlamydia sp.
 OS WO200208267-A2.
 PN 31-JAN-2002.
 PD 20-JUL-2001; 2001WO-US023121.
 XX 20-JUL-2000; 2000US-00620412.
 PR 23-APR-2001; 2001US-00841132.
 XX (CORI-) CORIXA CORP.
 PA Fling SP, Skeiky YAW, Probet P, Bhatia A;
 PI WPI; 2002-179901/23.
 XX Novel compositions comprising Chlamydia CapI protein and its use in the
 PT treatment of Chlamydia infection.
 XX Disclosure; Page 221-223; 537pp; English.

The present invention describes compositions comprising a Chlamydia CapI
 protein and methods for the diagnosis and therapy of Chlamydia infection.
 Chlamydia DNA and protein sequences from the present invention can have
 antibacterial and immunostimulant activities, and can be used in
 vaccines. Compounds from the present invention can be used for eliciting
 an immune response, specifically stimulating a Chlamydia-specific T-cell
 response or inhibiting the development of a Chlamydia infection in an
 animal. Methods from the present invention can be used: for detecting the
 presence of Chlamydia in a patient; to stimulate and/or expand T cells
 specific for a Chlamydia protein; and for treatment of a Chlamydia
 infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 sequences used in the exemplification of the present invention.

Sequence 982 AA;

Alignment Scores:
 Pred. No.: 0 Length: 982
 Score: 5084.00 Matches: 978
 Percent Similarity: 99.8% Conservative: 2

Best Local Similarity: 99.6% Mismatches: 2
 Query Match: 64.5% Indels: 0
 DB: 5 Gaps: 0
 US-10-701-844-1 (1-4435) x ABB94172 (1-982)
 QY 472 ATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGTATATCAATTTCCCTATATCTGTT 531
 Db 1 MetIleProGlnGlyIleTyArgGlyGluThrValIserPheProTyThrVal 20
 QY 532 ATAGAGATCCGAGTGGGACACTCTGTTTCTTCGAGGAGACTTAACATTAATAATCTT 591
 Db 21 IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuIysAsnLeu 40
 QY 592 GACAATTTCTATGACGCTTTGCTTTTAAAGTTCTTTTGGAACTTATTAGGAGTTTACT 651
 Db 41 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60
 QY 652 GTTTTAGGAGAGGACACTCGTTGACTTTCGAGACATACCGACTTCTACAAATGGGCA 711
 Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
 QY 712 GCTCTAAGTAATAGCGCTGCTGACTGCTTACTTATTGAGGGTTTAAAGAAATATCC 771
 Db 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheIysGluLeuSer 100
 QY 772 TTTTCCCAATTGCAATTCATTACTGCGTACTGCTGCTGCAACGACTTAATAGGCTAGC 831
 Db 101 PheSerAsnCyAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnIysGlySer 120
 QY 832 CAGACTCCGACGACCAACATCTACACCGTCTAAATGGTACTTATTTCTTAAACAGATCTT 891
 Db 121 GlnThrProThrThrSerThrProSerAsnGlyThrIleTySerIysThrAspLeu 140
 QY 892 TTGTTACTCAATAATGAGAGTTCATCTCTATAGTAATTTAGTCTCTCGAGATGGGGA 951
 Db 141 LeuLeuAsnAsnGluIysPheSerPheTySerAsnLeuValSerGlyAspGlyGly 160
 QY 952 GCTATAGATGCTAAGAGCTTAAACGGTTCAAGAAATTTAGCAAGCTTTGTCTTCCAGAA 1011
 Db 161 AlaIleAspAlaIysSerLeuThrValGlnGlyIleSerIysLeuCysValPheGlnGlu 180
 QY 1012 AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCT 1071
 Db 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200
 QY 1072 AACGAGCTCTATTGCTTTCCTGCGAATGTTGACGAGTAGAGAGGGGAGGATGCT 1131
 Db 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 220
 QY 1132 GCTGTTCCAGATGGGCGAGGAGTGTCATCTACTTCAACAGAAATCCAGTAGTA 1191
 Db 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAspProValVal 240
 QY 1192 AGTTTTTCCAGAAATACTCGGTTAGATTGTGGAACGTAGCCCGAGTAGGAGGAGG 1251
 Db 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260
 QY 1252 ATTTACTCTACGGGAACGTTGCTTTCCTGAATAATAGGAACCTTGTCTTCAACAT 1311
 Db 261 IleTySerTyArgAsnValAlaPheLeuAsnAsnGlyIysThrLeuPheLeuAsnAsn 280
 QY 1312 GTTGCTTCTCTGTTTACATCTGCTAAGCAACCAAGTCGACAGGCTTCTAATACG 1371
 Db 281 ValAlaSerProValTyrlleAlaAlaIysGlnProThrSerGlyGlnAlaSerAsnThr 300
 QY 1372 AGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAGAATGTCGCGAAGCAGGATCC 1431
 Db 301 SerAsnAsnTyArgAspGlyGlyAlaIlePheCysIysAsnGlyAlaGlnAlaGlySer 320
 QY 1432 AATAACTCTGGATCAGTTTCTCTTTGATGGAGAGGAGTAGTTTCTTTTAGTCAATGTA 1491
 Db 321 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 340

QY 1492 GCTGCTGGGAAGGGGAGCTATTATTCGCAAAAAGCTCTGGTGTCTAACTGGCCCT 1551
 Db 341 AlaaIaGlyLyseGlyAlaIaIeTyRAlaLyseLeuSerValAlaAsnCyseGlyPro 360
 QY 1552 GTACAAATTTAAAGGAATATCGTAATCATGGTGGAGGATTTATTTAGGAGAACTCTGA 1611
 Db 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyAlaIaIeTyRLeuGlyGluSerGly 380
 QY 1612 GAGCTCAGTTTCTGCTGATTATCGAGATATATTTTCGATGGGAATCTTAAAGAACA 1671
 Db 381 GluLeuSerLeuSerAlaAspTyRgLyAspIleIlePheAspGlyAsnLeuLyseArgThr 400
 QY 1672 GCCAAAGAAATGTCGCCGATGTTAATGGCGTAATCTGTGCTCAAGCCATTTCCGATG 1731
 Db 401 AlalyeGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420
 QY 1732 CGATCGGGAGGAAAATACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTCTTAAT 1791
 Db 421 GlySerGlyGlyLyseIleThrThrLeuArgAlaLyseAlaGlyHisGlnIleLeuPheAsn 440
 QY 1792 GATCCCATCGATGGCAACCGAAATAACACGAGCGAGCTCTTCCAAACTTCTAAAA 1851
 Db 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLyseLeuLyse 460
 QY 1852 ATTAACGATGGTAAGGATACACAGGGATATTGTTTTCGTAATGGAGCAATGACTTG 1911
 Db 461 IleAsnAspGlyGluGlyTyRThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
 QY 1912 TACCAAAATGTTACGATAGCAAGCAAGGATTTCTTCGTAAGCAAGCAAAATATCA 1971
 Db 481 TyRGINAsnValThrIleGlnGlyArgIleValLeuArgGluLyseAlaLyseLeuSer 500
 QY 1972 GTGAATCTCTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGAGTACATGGGAT 2031
 Db 501 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyRMetGluAlaGlySerThrLeuAsp 520
 QY 2032 TTTGTAACCTCCAAACCAACCAAGCTCTCGCGCTAATCAGTTGATCAGCTTTCC 2091
 Db 521 PheValThrProGlnProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer 540
 QY 2092 AATCTGCAATCTCTCTCTCTCTTTGTTAGCAAAATGCAATGCAATGCAATCTCTACC 2151
 Db 541 AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560
 QY 2152 AATCTCCAGCGCAAGATTCTCATCTCGAGTCATTGGTAGCACAACTGCTGTTCTGTT 2211
 Db 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrAlaGlySerVal 580
 QY 2212 ACAATTAGTGGCCCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGGTATGAT 2271
 Db 581 ThrIleSerGlyProfilePhePheGluAspLeuAspAspThrAlaTyRAspArgTyRAsp 600
 QY 2272 TGCTAGGTTCTAATCAAAAATCAATGCTCTGAAATTCAGTTAGGACTAAGCCCCA 2331
 Db 601 TipLeuGlySerAsnGlnLyseIleAsnValLeuLyseLeuGlnGlyThrLyseProPro 620
 QY 2332 GCTAATGCCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAGTATGCTATCAAGGA 2391
 Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLyseTyRgLyTyRglnGly 640
 QY 2392 AGCTGGAAGCTTCGCTGGGATCTTAATACAGCAAAATAAATGGTCTTATCTCTGAAGCT 2451
 Db 641 SerTrpLyseLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyRThrLeuLyseAla 660
 QY 2452 ACATGCACTAAACTGGGTATATCTCGGCTGAGGAGTAGCTCTTTGGTTCCAAAT 2511
 Db 661 ThrTrpThrLyseThrGlyTyRAsnProGlyProGluArgValAlaSerLeuValProAsn 680
 QY 2512 AGTTTATGGGATCCCAATTTAGATATACGATCTCGCGCATTCAGCAATTCAGCAAGTGTG 2571
 Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700

QY 2572 GATGGCGCTCTTATTTGTCGAGGATTATGGGTTTCTGGAGTTTCAAATTTCTTCTATCAT 2631
 Db 701 AspGlyArgSerTyRysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyRHis 720
 QY 2632 GACCGCATGCTTTAGGTACGGATATCGGTATATTAGTGGGGTTATTCCTTAGGAGCA 2691
 Db 721 AspArgAspAlaLeuGlyGlnGlyTyRArgTyRileSerGlyGlyTyRSerLeuGlyAla 740
 QY 2692 AACTCTACTTTGGATCATCATGATGTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751
 Db 741 AsnSerTyRAspGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
 QY 2752 AAAGATATGATGATGTCGTTCCAAATCATCATGCTGCATAGGATCGTTTATCTATCT 2811
 Db 761 LysAspTyRValValCyseArgSerAsnHisAlaCyseIleGlySerValTyRLeuSer 780
 QY 2812 ACCCAACAGCTTTATGTGATCTCTATTTGTTGGAGATGCGTTTATCCGTGCTAGTAC 2871
 Db 781 ThrGlnGlnAlaLeuCyseGlySerTyRLeuPheGlyAspAlaPheIleArgAlaSerTyR 800
 QY 2872 GGGTTTGGAAATCAGCATATCAAAACCTCATATATCATTTGAGAGGAGCGATTTGCT 2931
 Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyRThrPheAlaGluGluSerAspValArg 820
 QY 2932 TGGGATATAACTCTCTGCTGAGAGATTCGAGCGGATACCGATTTGATGTTACTTCCA 2991
 Db 821 TrpAspAsnAsnCyseLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
 QY 2992 TCTAAGCTCTATTGAAATGAGTTGCTCTTCTGTCAGCTGAGTTTCTTATCCGAT 3051
 Db 841 SerLyseLeuTyRLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyRAlaAsp 860
 QY 3052 CATGAATCTTTTACAGAGAAAGCGATCAAGCTGGGCAATTCAGAGCGGACATCTCTA 3111
 Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
 QY 3112 AATCTATCAGTTCTCTGTTGGAGTGAAGTTGATCGATGCTTAGTACACATCTTAATAA 3171
 Db 881 AsnLeuSerValProValGlyValLysePheAspArgCyseSerSerThrHisProAsnLyse 900
 QY 3172 TATAGCTTTATGGCGCTTATATCTGTGATGCTTATTCGACCATCTCTGGTACTGAGACA 3231
 Db 901 TyRSerPheMetAlaAlaTyRileCyseAspAlaTyRArgThrIleSerGlyThrGluThr 920
 QY 3232 AGCTCTCTATCCCATCAAGAGACATGAGACACAGATGCTTTCATTTAGCAAGACATGA 3291
 Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940
 QY 3292 GTTGTGGTTAGGATCTATGATGCTTCTCTTAAAGTAAATATAGATATATGGCCAT 3351
 Db 941 ValValValArgGlySerMetTyRAlaSerLeuThrSerAsnIleGluValTyRgLyHis 960
 QY 3352 GGAAGATATGATGATGATGATGCTTCTGAGGCTATGTTGAGTGCAGGAAAGTAGAGTC 3411
 Db 961 GlyArgTyRgLyTyRArgAspAlaSerArgGlyTyRgLyLeuSerAlaGlySerLyseVal 980
 QY 3412 CGGTTT 3417
 Db 981 ***Phe 982
 RESULT 29
 AAB13639
 ID AAB13639 standard; protein; 1006 AA.
 XX
 AC AAB13639;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE C. trachomatis pmpG gene protein.
 XX
 KW Chlamydial infection; sexually transmitted disease;
 KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
 KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;

KW coronary heart disease; antibacterial.
 XX Chlamydia trachomatis.
 OS WO200034483-A2.
 XX 15-JUN-2000.
 XX 08-DEC-1999; 99WO-US029012.
 XX 08-DEC-1998; 98US-00208277.
 XX 08-APR-1999; 99US-00288594.
 PR 01-OCT-1999; 99US-00410568.
 PR 22-OCT-1999; 99US-00426571.
 XX (CORI-) CORIXA CORP.
 PA Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
 PI WPI; 2000-431303/37.
 XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
 PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
 PT acid sequence encoded by polynucleotide sequence.
 XX Claim 2; Page 208-210; 256pp; English.
 PS The present invention relates to new nucleic acid sequences and the
 XX proteins encoded by the nucleic acid sequences. The encoded proteins
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia
 CC infection. Chlamydiae are intracellular bacterial pathogens that are
 CC responsible for a wide variety of human infections. C. trachomatis
 CC infection is one of the most common sexually transmitted diseases and can
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a
 CC major cause of acute respiratory tract infections in humans and is also
 CC thought to play a role in the pathogenesis of atherosclerosis and
 CC coronary heart disease. The present sequence is a protein isolated in the
 CC present invention
 XX Sequence 1006 AA;
 SQ

Alignment Scores:
 Pred. No.: 0 Length: 1006
 Score: 5083.00 Matches: 978
 Percent Similarity: 99.8% Conservative: 2
 Best Local Similarity: 99.6% Mismatches: 2
 Query Match: 64.5% Indels: 0
 DB: 3 Gaps: 0

US-10-701-844-1 (1-4435) x AAB13639 (1-1006)

QY 472 ATGGTCTCTCAAGAAATACGATGGGAGACGTTAACTGATCATTTCCCTACTACTGTT 531
 DB 25 MetileProGinglyIleThyAspGlyGluThrLeuThrValSerPheProThyThrVal 44
 QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCTCTCGAGAGAGTTAACTTAAATAATCTTT 591
 DB 45 IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64
 QY 592 GACAAATTTCTATGAGCTTTGCTTAAAGTTGTTTGGGAACCTTATTAGGAGTTTACT 651
 DB 65 AspAsnSerIleAlaLeuProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 84
 QY 652 GTTTTAGGAGAGACACTGGTTCATCTTCGAGACATACGACTTCTCAAAATGGGCA 711
 DB 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104
 QY 712 GCTCTAAGTAATACGCTGCTGATGAGTCTGTTTACTATTGAGGCTTTTAAAGAAATATCC 771
 DB 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124

QY 772 TTTTCAATTGCAATTCATTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGTGAC 831
 DB 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaIleThrThrAsnLysGlySer 144
 QY 832 CAGACTCCGAGACAAACATCTACACCCGCTTAATGCTACTATTATTCTTAAACACATCTT 891
 DB 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleThySerLysThrAspLeu 164
 QY 892 TTGTTACTCAATATGAGAAGTTCTCATTCTATAGTAATTTAGTCTCTCGAGATGGGGA 951
 DB 165 LeuLeuLeuAsnAsnGlnLysPheSerPheThySerAsnLeuValSerGlyAspGlyGly 184
 QY 952 GCTATAGATGCTAAGAGCTTAAACGTTTCAAGGAATTAGCAAGCTTTGCTCTTCCAAGAA 1011
 DB 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 204
 QY 1012 AATAGTCTCAAGCTGATGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCT 1071
 DB 205 AsnThrAlaGlnAlaAspGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224
 QY 1072 AACGAGGCTCCTATTGCTTTGTAGCGAATGTTGAGGAGTAGTAAGAGGGGAGGATTGCT 1131
 DB 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244
 QY 1132 GCTGTTTCAAGATGGCAGGAGTGTCTCATCTTCAACAGAAAGATCCAGTAGTA 1191
 DB 245 AlaValGlnAspGlyGlnGlyValSerSerSerThrThrGluAspProValVal 264
 QY 1192 AGTTTTTCCAGAAATCTCCGCTAGATTGATGGAGAGTAGCCGAGTAGGAGGAGG 1251
 DB 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 284
 QY 1252 ATTTTACTCTACGGAGCGTCTCTTCTCTGAATAATGGAATAACCTTGTCTTCAACAAT 1311
 DB 285 IleThySerThyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 304
 QY 1312 GTTGCTTCTCTGTTTACATTTCTGCTAAGCAACCAACAGTGGACAGCTTCTAATACG 1371
 DB 305 ValAlaSerProValThyIleAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324
 QY 1372 AGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAGATGGTGGCAGAGAGATCC 1431
 DB 325 SerAsnAsnThyGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344
 QY 1432 AATACTCTGGATCAGTTTCTCTTGTGAGAGGAGTAGTTTCTTATAGTCAATGTA 1491
 DB 345 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 364
 QY 1492 GCTGCTGGGAAAGGGGAGCTATTTATGCAAAAAGCTCTCGGTTGCTAATCTGCCCCCT 1551
 DB 365 AlaAlaGlyLysGlyAlaIleThyAlaLysLysLeuSerValAlaAsnCysGlyPro 384
 QY 1552 GTACAAATTTTAAAGGAATATCGTAATGATGGTGGAGCGCATTTATTAGAGAACTCGGA 1611
 DB 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleThyLeuGlyGluSerGly 404
 QY 1612 GAGCTCAGTTTATCTCTGATTTATGAGATATTATTTTCGATGGGAATCTTAAAGAACAA 1671
 DB 405 GluLeuSerLeuSerAlaAspThyGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424
 QY 1672 GCCAAAGAGAATGCTCCGATGTTAATGGCTAATGCTGTCTCTCACAGCCATTTCCGATG 1731
 DB 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444
 QY 1732 GGATCGGGAGGGAAAAATAACGACATTAAAGAGCTTAAAGAGGGCATCATGATCTCTTAAT 1791
 DB 445 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464
 QY 1792 GATCCCATCGAGATGCCAAACCGAATTAACCCAGCGCAGCTTCCCAACTTCTTAAAA 1851
 DB 465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484

QY 1852 ATTAACGATGGTGAAGGATACACAGGGATATGTTTTGCTTAATGGAGCAGTACTTTG 1911
 Db 485 TleAenAepGlyGluGlyThrGlyAepIleValPheAlaAenGlySerSerThrLeu 504
 QY 1912 TACCAAAATGTTACCATAGACAGGAAGGATGTTCTTCTGTAAGAAAGCAAAATATCA 1971
 Db 505 TyrGlnAenValThrIleGluGlnIleValLeuArgGluValAlaAenLeuSer 524
 QY 1972 GTGAATCTCTTAAGTCAGACAGGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGAT 2031
 Db 525 ValAenSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAep 544
 QY 2032 TTTGTAACTCCAAACCCACCAACAGCTCTCGCGCTTAATCATGATTGATCAGCTTTTCC 2091
 Db 545 PheValThrProGlnProGlnProGlnProAlaAlaAenGlnLeuIleThrLeuSer 564
 QY 2092 AATCTGCATTTGCT 2151
 Db 565 AenLeuHleLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThrAenProProThr 584
 QY 2152 AATCTCCAGCGCAAGATTCTCATCTGCAGTCATTGGTAGCACAACTGCTGTTCTGT 2211
 Db 585 AenProProAlaGlnAepSerHisProAlaValIleGlySerThrAlaGlySerVal 604
 QY 2212 ACAATTAGTGGCCCTATCTTTTTCAGGATTTGGATGATAGACCTTATGATAGTATGAT 2271
 Db 605 ThrIleSerGlyProIlePheGluAepLeuAepAepThrAlaTyrAepAepGlyrAep 624
 QY 2272 TGGCTAGGTTCTTAATCAAAATCAATGCTCTGAAATACAGTATAGGACCTAAGCCCA 2331
 Db 625 TrpLeuGlySerAenGlyLysIleAenValLeuLysLeuGlnLeuGlyThrLysProPro 644
 QY 2332 GCTAATGCCCATCAGATTGATCTCTAGGGAATGAGATGCCCTTAAGTATGCTATCAAGA 2391
 Db 645 AlaAenAlaProSerAepLeuThrLeuGlyAenGluMetProLysTyrGlyThrGlnGly 664
 QY 2392 AGCTGGAAGCTTGGTGGATCCTTAATACAGCAATTAATGGTCTTATCTCTGAAGCT 2451
 Db 665 SerTrpLysLeuAlaTrpAepProAenThrAlaAenAenGlyProTyrThrLeuLysAla 684
 QY 2452 ACATCGACTAAACCTGGGTATATCTGGGCTGAGCAGTAGTACTCTTTGGTTCCTCAAT 2511
 Db 685 ThrTrpThrLysThrGlyTyrAenProGlyProGluArgValAlaSerLeuValProAen 704
 QY 2512 AGTTATGGGGATCCATTTTAGATATACGATTCGCGCATTCAGCAATTCAGCAAGTGTG 2571
 Db 705 SerLeuTrpGlySerIleLeuAepIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
 QY 2572 GATGGGCGCTTATTTGTCGAGATTAATGGGTTCTGGAGTTTCGAAATTTCTTCTATCAT 2631
 Db 725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAenPhePheTyrHis 744
 QY 2632 GACCGCATGCTTTAGTTCAGGATATCGTATATTTAGTGGGGTATTCTTCTTAGGACA 2691
 Db 745 AspaAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
 QY 2692 AACTCTACTTTGATTCATGATGTTGCTCTAGCATTTTACCAAGTATTGCTAGATCT 2751
 Db 765 AenSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
 QY 2752 AAAGATTATGATGTGTGCTTCCCAATCATCATGCTTGGATAGGATCCGTTTATCTATCT 2811
 Db 785 LysAepTyrValValCysArgSerAenHisAlaCysIleGlySerValTyrLeuSer 804
 QY 2812 ACCCAACAGCTTATGTCGATCTATTTGTCGAGATGCGTTTATCGTCTAGCTAC 2871
 Db 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAepAlaPheIleArgAlaSerTyr 824
 QY 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTTGCAGAGGAGCGATGTTCT 2931
 Db 825 GlyPheGlnAenGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAepValArg 844
 QY 2932 TGGGATATTAATGCTCTGCTCGAGAGATGGAGCGGATACCGATTGCTGATTACTCCA 2991

Db 845 TrpAenAenCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864
 QY 2992 TCTAAGCTCTATTGAATAGTTCGCTCTTTCGTCGAAAGCTGAGTTTCTTATGCCGAT 3051
 Db 865 SerLysLeuTyrLeuAenGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAep 884
 QY 3052 CATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGCAATTCAGAGCGGACATCTCCTA 3111
 Db 885 HisGluSerPheThrGluGluGlyAepGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
 QY 3112 AATCTCATCAGTCTCTGTTGGAGTGAAGTTTCATCGATGCTTCTAGTACACATCTTAATAA 3171
 Db 905 AenLeuSerValProValGlyValLysPheAepArgCysSerSerThrHisProAenLys 924
 QY 3172 TATAGCTTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231
 Db 925 TyrSerPheMetAlaAlaTyrIleCysAepAlaTyrArgThrIleSerGlyThrGluThr 944
 QY 3232 AGCTCTCTATCCCATCAAGACACATGGAACAACAGATGCTTTCATTTTAGCAAGACATGA 3291
 Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAepAlaPheHisLeuAlaAargHisGly 964
 QY 3292 GTTGTGGTTAGAGGATCTATGATGCTTCTCTAAACAAGTAATATAGAAGTATATGCCAT 3351
 Db 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAenIleGluValTyrGlyHis 984
 QY 3352 GGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGGAGTCGAGGAGTAGAGTC 3411
 Db 985 GlyArgTyrGluTyrArgAepAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004
 QY 3412 CGGTTTC 3417
 Db 1005 ArgPhe 1006

RESULT 30
 ADD42756
 ID ADD42756 standard; protein; 670 AA.
 XX AC ADD42756;
 XX DT 15-JAN-2004 (first entry)
 XX DE Chlamydia pmpG passenger domain protein SEQ ID NO:169.
 XX KW Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
 KW antileishmaniasis; cardiant; antiarteriosclerotic; ophthalmological;
 KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
 KW tubal obstruction; infertility; male infertility; ocular infection;
 KW blindness; acute respiratory tract infection; atherosclerosis;
 KW coronary heart disease.
 XX OS Chlamydia sp.
 XX PN WO2003041560-A2.
 XX PD 22-MAY-2003.
 XX PF 05-NOV-2002; 2002WO-US035624.
 XX PR 06-NOV-2001; 2001US-00012256.
 PR 05-DEC-2001; 2001US-00007693.
 PR 15-JUL-2002; 2002US-00197220.
 XX (CORI-) CORIXA CORP.
 XX Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;
 PI Probst P;
 XX WPI; 2003-441771/41.
 DR N-PSDB; ADD42747.
 XX PT New DNA and proteins comprising a portion of a Chlamydia antigen, useful

PT for diagnosing or treating Chlamydial infections, particularly as
 PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
 XX inflammatory disease.
 XX
 XX Claim 2; SEQ ID NO 169; 275pp; English.
 CC The present invention describes compounds and methods for diagnosing and
 CC treating Chlamydial infection. Chlamydia polynucleotide and protein
 CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
 CC antiarteriosclerotic and ophthalmological activities, and can be used in
 CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
 CC compositions or methods from the present invention can be used for the
 CC serodiagnosis or treatment of Chlamydial infections, particularly in
 CC humans. The polynucleotides, proteins or compositions are particularly
 CC useful for stimulating an immune response in a patient, or for
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC Specifically, the polynucleotides, proteins or compositions are useful as
 CC vaccines for treating or preventing Chlamydial infections including
 CC pelvic inflammatory disease (which results in tubal obstruction and
 CC infertility in women), male infertility, ocular infection (which may
 CC cause blindness), acute respiratory tract infections, atherosclerosis, or
 CC coronary heart disease. The present sequence is used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 670 AA;

Alignment Scores:
 Pred. No.: 5,848-289 Length: 670
 Score: 3336.50 Matches: 651
 Percent Similarity: 98.8% Conservative: 11
 Best Local Similarity: 97.2% Mismatches: 7
 Query Match: 42.3% Indels: 1
 DB: 7 Gaps: 1

US-10-701-844-1 (1-4435) x ADD42756 (1-670)

QY	463	CGAGAAATCATGCTTCTCAAGAAATTTACGATGGGGAGACGTTAACTGTATCTTCC	522
DB	1	AlaGluIleMeIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPhePro	20
QY	523	TATACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTCTCGAGGAGGTTAACTTA	582
DB	21	TyrThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu	40
QY	583	AAAAATCTTGACAAATCTATTGAGCTTTGCCCTTTAAGTGTGTTGGGAACTATTAGGG	642
DB	41	LysAsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly	60
QY	643	AGTTTACTGTTTATAGGAGGACACTCGTTGACTTTTCGAGAACATACGACTTCTACA	702
DB	61	SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr	80
QY	703	AATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTGAGGGTTTAAA	762
DB	81	AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys	100
QY	763	GAAATATCTTTTCCAAATTCGAAATTCATTATTCGCGTACTGCTGCTGCAACGACTAAT	822
DB	101	GluLeuSerPheSerAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsn	120
QY	823	AGGGTAGCCAGACTCCGACGACCAACATCTACACCGCTTAATGCTACTATTATCTAAA	882
DB	121	AsnGlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleYrSerLys	140
QY	883	ACAGATCTTTTGTACTCAATAATGAGAACTTCTCATTTCTATAGTAAATTTAGTCTCTGA	942
DB	141	ThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPheYrSerAsnLeuValSerGly	160
QY	943	GATGGGGAGCTATAGAGCTTAAGAGCTTAACCGTTCAAGGAATTAGAGCTTTGTGTC	1002
DB	161	AspGlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysVal	180
QY	1003	TTCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAGAGTAGTCACCAGTTTCTCT	1062

181	PheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSer	200
1063	GCTATGGCTAACGAGGCTCTATTGCTTTCCTAGCAATGTTGCGAGGAGTAGAGGGGA	1122
201	AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly	220
1123	GGGATTGCTGCTGTTTACGAGTGGGAGGAGTGTCTCATCTTCAACAGAGAT	1182
221	GlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAsp	240
1193	CCAGTAGTAAGTTTTCCAGAAATCTCGGTAGAGTTTGTGGAAGAGTACCGGCTA	1242
241	ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal	260
1243	GGAGGAGGATTACTCTCTACGGGAACGTTCTCTCTGAATAATGGAATAACCTTGT	1302
261	GlyGlyGlyIleYrSerYrGlyAsnValAlaPheLeuAsnAsnGlyYrThrLeuPhe	280
1303	CTCAACAATGTTGCTTCTCTGTTTACATTTGCTGCTAAGCAACCAAGTCGACAGGT	1362
281	LeuAsnAsnValAlaSerProValTyrIleAlaGluGlnProThrAsnGlyGlnAla	300
1363	TCTAATACGATTAATAATACGAGATGGAGAGCTATCTTCTGTAAGNATGTCGCA	1422
301	SerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGln	320
1423	---GCAGGATCCAAATACTCTGGATCAGTTTCTTGTGATGGAGGAGTAGTTTCTTT	1479
321	AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePhe	340
1480	AGTAGCAATGATGCTGCTGGGAAAGGGGAGCTATTATTATGCCAAAAAGCTCTCGT	1539
341	SerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAla	360
1540	AACTGCGCCCTGTACAATTTTAAAGGAATATCGCTAATATGATGGTGGACGATTTATTA	1599
361	AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyrLeu	380
1600	GGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGAGATATATTATTTCCATGGAAAT	1659
381	GlyGluSerGlyGluLeuSerLeuSerAlaAspYrGlyAspIleIlePheAspGlyAsn	400
1660	CTTAAAAAGAACCAAGAGAAATCTGCCGATGTTAATGGCGTAACTGTGTCTCACAA	1719
401	LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln	420
1720	GCCATTTCCGATCGGATCGGGAGGAAATAACGACATTAAGAGCTAAAGCAGGCGATCAG	1779
421	AlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln	440
1780	ATTCTCTTTAATGATCCCATCCAGATGGCAACGGGAAATAACAGCCAGCGAGTCTTCC	1839
441	IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer	460
1840	AAACTTCTAAAAATTAACGATGCTGAAGGATACACAGGGGATATGTTTTGCTAATGGA	1899
461	GluProLeuLysIleAsnAspGlyGluGlyYrThrGlyAspIleValPheAlaAsnGly	480
1900	AGCAGTACTTGTACCAAAATGTTACGATAGACGAAGGAGGATTTCTTCTGTTGAAAAG	1959
481	AsnSerThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLys	500
1960	GCAAAATATCAGTGAATTTCTTAAGTCAGACAGGTGGGAGCTGTATATGAGAGCTGG	2019
501	AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuYrMetGluAlaGly	520
2020	AGTACATGGGATTTGTAACTCCACAAACCAACCAACAGCTCTCTCCGCTAATCAGTTG	2079
521	SerThrLeuAspPheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeu	540
2080	ATCACGCTTTTCCAATCTGCAATTTGCTCTTCTTTCTTTGTAGCAAAACATGAGTTACG	2139

Db 541 IleThrLeuSerAenLeuHisLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThr 560
 QY 2140 AATCCTCTACCAATCTCTCAGCGCAAGATTCTCATCTCTGAGTCATTGGTAGCACAACT 2199
 Db 561 AenProThrAenProAlaGlnAenSerHisProAlaIleIleGlySerThrThr 580
 QY 2200 GCTGGTCTGTTTACAAATAGTGGCCTATCTTTTGGAGATTGGATGATACAGCTTAT 2259
 Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAenLeuAenAenThrAlaTyr 600
 QY 2260 GATAGGTATGATGGCTAGGTCTTAATCAAAAAATCAATGCTCCTGAAATACAGTTAGG 2319
 Db 601 AspArgTyrAenTyrLeuGlySerAenGlnIleAenValLeuLeuGlnLeuGly 620
 QY 2320 ACTAAGCCCCCAGCTAATGCCCCATCAGATTGACTCTAGGAAATGAGATGCCCTAAGTAT 2379
 Db 621 ThrGlnProSerAlaAenAlaProSerAenLeuThrLeuGlyAenGluMetProIleTyr 640
 QY 2380 GCGTATCAAGGAAGCTGGAAGCTTGGTGGATCTTAATACAGCAAAATTAATGGTCTTAT 2439
 Db 641 GlyTyrGlnGlySerTyrLeuAlaTyrAenProAenThrAlaAenAenGlyProTyr 660
 QY 2440 ACTGTAAAGCTACATGCACTAAACTGGG 2469
 Db 661 ThrLeuLysAlaThrTyrThrLysThrGly 670

RESULT 31

ID AAY37238 standard; protein; 524 AA.

XX AAY37238;
 XX
 XX
 DT 07-OCT-1999 (first entry)
 XX

Chlamydia trachomatis cellular envelope protein.

Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
 nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 bartholinitis; pneumopathy; venereal lymphogranulomatosis.

Chlamydia trachomatis.

W09928475-A2.

10-JUN-1999.

27-NOV-1998; 98WO-IB001939.

28-NOV-1997; 97FR-00015041.

17-DEC-1997; 97FR-00016034.

04-NOV-1998; 98US-0107077P.

(GEST) GENSET.

Griffais R;

WPI; 1999-371125/31.

Genome sequence of Chlamydia trachomatis.

Disclosure; Page 996-997; 1755pp; English.

AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
 of Chlamydia trachomatis (see AAY01425). The polypeptides can be used as
 vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 can also be used to control growth of the microorganism. Chlamydia
 trachomatis is responsible for a large number of diseases, e.g. eye
 diseases such as conventional trachoma, nonendemic trachoma,
 paratrachoma, and inclusion conjunctivitis; genital diseases such as
 nongonococcal urethritis, epididymitis, cervicitis, salpingitis, and
 perihhepatitis, bartholinitis; pneumopathy in breast feeding infants; and
 venereal lymphogranulomatosis. The polypeptides of the invention may be

CC of use in treating these diseases
 XX
 SQ Sequence 524 AA;
 Alignment Scores:
 Pred. No.: 1.3e-223 Length: 524
 Score: 2606.00 Matches: 518
 Percent Similarity: 99.2% Conservative: 0
 Best Local Similarity: 99.2% Mismatches: 4
 Query Match: 33.1% Indels: 1
 Gaps: 0

US-10-701-844-1 (1-4435) x AAY37238 (1-524)

QY 625 TTTGGGAACCTATTAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTTCGAG 684
 Db 3 PheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPheGlu 22
 QY 685 AACATACGACCTTCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTT 744
 Db 23 AenIleArgThrSerThrAenGlyAlaAlaLeuSerAenSerAlaAlaAenGlyLeuPhe 42
 QY 745 ACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTACTTCCGCTACTG 804
 Db 43 ThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaValLeu 62
 QY 805 CCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACCGTCTAAT 864
 Db 63 ProAlaAlaThrThrAenLysGlySerGlnThrProThrThrSerThrProSerAen 82
 QY 865 GGTACTATTATCTTAAACAGATCTTTTGTGTACTCAATATAGAGATTTCTCATTTCTAT 924
 Db 83 GlyThrIleTyrSerLysThrAspLeuLeuLeuAenGluLysPheSerPheTyr 102
 QY 925 AGTATTATTAGTCTCGGAGATGGGGAGCTATAGATCTAGAGCTTAACGGTTCAAGCA 984
 Db 103 SerAenLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGlnGly 122
 QY 985 ATTAGCAAGCTTTGTGCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAA 1044
 Db 123 IleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAenGlyGlyAlaCysGln 142
 QY 1045 GTAGTCAACAGTTTCTGCTATGCTTAACGAGGCTCTATTGCTTTGTAGCGAATGTT 1104
 Db 143 ValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheValAlaAenVal 162
 QY 1105 GCAGAGTAAAGGGGGAGGATGCTGCTGTTCAAGATGGGACAGAGGAGTGTCACTCA 1164
 Db 163 AlaGlyValArgGlyGlyIleAlaAlaValGlnAenGlyGlnGlnGlyValSerSer 182
 QY 1165 TCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTTCCAGAAATACCTGCGTAGAGTTGAT 1224
 Db 183 SerThrSerThrGluAenProValValSerPheSerArgAenThrAlaValGluPheAen 202
 QY 1225 GGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTTCTCGAAT 1284
 Db 203 GlyAenValAlaArgValGlyIleTyrSerTyrGlyAenValAlaPheLeuAen 222
 QY 1285 AATGGAAAAAAGCTTTGTTTCTCAACAAATGTTGCTTCTCTGTTTACATTTGCTGAAGCA 1344
 Db 223 AenGlyLysThrLeuPheLeuAenAenValAlaPheProValTyrIleAlaAlaLysGln 242
 QY 1345 CCAACCAAGTGAAGAGCTTCTTAATACGAGTAGTAATTAATACGGAGATGGAGGAGCTATCTTC 1404
 Db 243 ProThrSerGlyGlnAlaSerAenThrSerAenAenTyrGlyAspGlyGlyAlaIlePhe 262
 QY 1405 TGTAAGAAATGTCGCAAGAGGAGTCCAAATACCTCTGGATCAGTTTCTTCTGATGAGAG 1464
 Db 263 CysLysAenGlyAlaGlnValGlySerAenAenSerGlySerValSerPheAenGlyGlu 282
 QY 1465 GGAGTAGTTTCTTTTATAGCAATGCTAGCTGCTGGAAAGGGGAGCTATTTATGCAAAA 1524
 Db 283 GlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLys 302


```
Db 181 GlnGluAenThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAla 200
Qy 1066 ATGGCTAACGAGGCTCTTATTGCTTCTAGCAATGTTGCAGGAGTAAAGGGGAGGG 1125
Db 201 MetAlaAenGluAlaProIleAlaPheValAlaAenValAlaGlyValArgGlyGly 220
Qy 1126 ATTGCTGCTTTCAGGATGGGAGGAGGAGTGCATCATCTTCAACAGAGATCCA 1185
Db 221 IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspPro 240
Qy 1196 CTAGTAAGCTTTTCCAGAAATCTCGCGTGAAGTGTGGAACGTAAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerArgAenThrAlaValGluPheAspGlyAenValAlaArgValGly 260
Qy 1246 GGAGGATTTACTCTCAGGGAACCTGCTTCTCGAATAATGGAATAACCTTGTTCCTC 1305
Db 261 GlyGlyIleTyrSerTyrGlyAenValAlaPheLeuAenAenGlyLysThrLeuPheLeu 280
Qy 1306 AACAAATGTTCTTCTCTGTTTACATTCGCTGAAGCAACCAAGTGGACAGCTTCT 1365
Db 281 AenAenValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
Qy 1366 AATACGACTAATAATACGAGATGGAGAGCTATCTCTGTAAGTAAAGTGGTGGCAAGCA 1425
Db 301 AenThrSerAenAenTyrGlyAspGlyGlyAlaIlePheCysLysAenGlyAlaGlnAla 320
Qy 1426 GGATCCAAATACCTGGAATCAGTTCTCTTGTGAGGAGGAGTAGTTCTTTAGTAGC 1485
Db 321 GlySerAenAenSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSer 340
Qy 1486 AATGTAGCTGTGGAAAGGGAGCTATTATATGCCCCAAAGCTCTCGTGTCTAACTGT 1545
Db 341 AenValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAenCys 360
Qy 1546 GGCCTGTACATTTTAAAGTAATCCCTAATGATGGTGGAGCATTTATTAGGAA 1605
Db 361 GlyProValGlnPheLeuArgAenIleAlaAenAspGlyGlyAlaIleTyrLeuGlyGlu 380
Qy 1606 TCTGAGAGCTCAGTTTATCTCTGATATGAGATATTATTTTCGATGGGAATCTTAAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAenLeuLys 400
Qy 1666 AGAACAGCCAAAGAGATGCTCCGATGTTAATGGCTAACTGTGTCTCACAAGCCATT 1725
Db 401 ArgThrAlaLysGluAenAlaAlaAspValAenGlyValThrValSerSerGlnAlaIle 420
Qy 1726 TCGATGGGATCGGAGGGGAAATACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
Qy 1786 TTTAATGATCCCATCGAGATGGCAAGGAAATACACGACGCGAGTCTTCCAAACTT 1845
Db 441 PheAenAspProIleGluMetAlaAenGlyAenAenGlnProAlaGlnSerSerLysLeu 460
Qy 1846 CTPAAATTAACGATGGTGAAGATACACAGGGGATTTGTTTTTGTCTAATGGAAGCAGT 1905
Db 461 LeuLysIleAenAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAenGlySerSer 480
Qy 1906 ACTTTGTACAAATGTTACGATAGCAGGAGGAGGATTTCTTCGTGCAAAAGCCAAA 1965
Db 481 ThrLeuTyrGlnAenValThrIleGluGlnGlyArgIleValLeuArgGlyLysAlaLys 500
Qy 1966 TTATCAGTGAATTCT 1980
Db 501 LeuSerValAspSer 505
```

RESULT 33

AAV16751

ID AAV16751 standard; protein; 158 AA.

XX

AC AAV16751;

XX

```
DT 21-JUL-1999 (first entry)
XX Chlamydia HMW protein fragment.
DS
XX Chlamydia; high molecular weight protein; HMW protein; urethritis;
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
KW salpingitis; tubal occlusion; infertility; cervical cancer;
KW arteriosclerosis; atherosclerosis.
OS Chlamydia sp.
XX
XX W09917741-A1.
XX 15-APR-1999.
XX 01-OCT-1998; 98WO-US020737.
XX 02-OCT-1997; 97US-00942596.
XX (ANTE-) ANTEX BIOLOGICS INC.
XX Jackson JW, Pace JL;
XX WPI; 1999-287659/24.
XX
XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
XX cervical cancer.
XX Claim 5; Page 133-135; 141pp; English.
XX
XX The invention relates to an isolated Chlamydia species high molecular
XX weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
XX determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
XX be used for preventing, treating or ameliorating a disorder related to
XX Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
XX lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
XX pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
XX cancer, infertility, arteriosclerosis and atherosclerosis. The products
XX can also be used for detection and diagnosis. Sequences AAY16740-Y16752
XX represent Chlamydia HMW protein fragments
XX
XX SQ Sequence 458 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,01e-200 Length: 458
XX Score: 2350.00 Matches: 458
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 29.8% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-10-701-844-1 (1-4435) x AAY16751 (1-458)
Qy 1030 GGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGCTATGGCTTAACGAGCTCTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAla 20
Qy 1090 TTTGTAGCGAATGTTGCAGGAGTAAAGGGGAGGATTCGCTGTTCCAGGATGGCAG 1149
Db 21 PheValAlaAenValAlaGlyValArgGlyGlyIleAlaValGlnAspGlyGln 40
Qy 1150 CAGGAGGTGTCATCATCTACTTCAACAGAAAGTCCAGTAGTAGTTTTCAGAAATACT 1209
Db 41 GlnGlyValSerSerSerThrThrGluAspProValValSerPheSerArgAenThr 60
Qy 1210 GCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAC 1269
Db 61 AlaValGluPheAspGlyAenValAlaArgValGlyGlyIleTyrSerTyrGlyAen 80
Qy 1270 GTTGTCTTCTCAATAATGGAATAACCTTGTTCCTCAACAATGTTGCTTCTCTGTTTAC 1329
Db 81 ValAlaPheLeuAenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyr 100
```

QY	1330	ATTGCTGCTAGCAACCAACGACGAGCTTCTAATACAGTAATAATACGAGAT	1389	RESULT 34	AAV16752	
DB	101	IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp	120	ID	AAV16752 standard; protein; 325 AA.	
QY	1390	GGAGGAGCTATCTCTGAAGAATGCTGGCAAGCAGGATCAATAACTCTCGATCAGTT	1449	AC	AAV16752;	
DB	121	GlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal	140	DT	21-JUL-1999 (first entry)	
QY	1450	TCCTTTGATGGAGGAGTACTTTTCTTTAGTACCAATGTAGCTGCTGGGAAAGGGA	1509	DE	Chlamydia HMW protein fragment.	
DB	141	SerPheAspGlyGluGlyValPhePheSerSerAsnValAlaAlaGlyGlyGly	160	XX	Chlamydia; high molecular weight protein; HMW protein; urethritis;	
QY	1510	GCTATTATGCCAAAAGCTCGGTTGCTTAAGTGGCCCTGTAACAATTTTAAAGGAT	1569	KW	bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;	
DB	161	AlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn	180	KW	cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;	
QY	1570	ATCGCTAAATGATGCTGGAGCGATTTATTTAGGAGATCTGGAGAGCTCAATTTATCTGCT	1629	XX	salpingitis; tubal occlusion; infertility; cervical cancer;	
DB	181	IleAlaAsnAspGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAla	200	XX	arteriosclerosis; atherosclerosis.	
QY	1630	GATTATGGAGATATTTTTCGATGGGAATCTTTAAAGAACAGCCAAAGAGATGCTGCC	1689	OS	Chlamydia sp.	
DB	201	AspTyrGlyAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla	220	PN	WO9917741-Al.	
QY	1690	GATGTTAATGGGTAACTGTGCTCTCAAGCCATTTTCATGGGATCGGAGGGAATA	1749	PD	15-APR-1999.	
DB	221	AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIle	240	PF	01-OCT-1998; 98WO-US020737.	
QY	1750	ACGACATTAAGAGTAAGCAGGCGATCTCTTAATGATCCATCGAGATGGCA	1809	PR	02-OCT-1997; 97US-00942596.	
DB	241	ThrThrLeuArgAlaLysAlaGlyHisGlnIlePheAsnAspProIleGluMetAla	260	XX	(ANTE-) ANTEX BIOLOGICS INC.	
QY	1810	AACGGAATAAACAGCAGCGAGTCTTCCAACTTCTAAAAATTAAACGATGCTGAAGGA	1869	PI	Jackson JW, Pace JL;	
DB	261	AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLysIleAsnAspGlyGluGly	280	XX	WPI; 1999-287659/24.	
QY	1870	TACACAGGGATATGTTTGTCTTAATGAGCAGTACTTTGTACCAAAATGTAGATA	1929	DR	New Chlamydia protein useful for treating conjunctivitis, urethritis and	
DB	281	TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle	300	PT	cervical cancer.	
QY	1930	GAGCAAGGAGGATGTTCTTCTGAAAGGCAAAATATACGTGATTTCTTAAGTCAG	1989	XX	Claim 5; Page 135-136; 141pp; English.	
DB	301	GluGlnGlyArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGln	320	CC	The invention relates to an isolated Chlamydia species high molecular	
QY	1990	ACAGTGGGAGTCTGTATATGGAAGCTGGGAGTACATGGATTTTGTAACTCCACACCA	2049	CC	weight (HMW) protein having an apparent mol. wt. of 105-115 kD as	
DB	321	ThrGlyGlySerLeuTyrMetGluAlaGlySerThrTrpAspPheValThrProGlnPro	340	CC	determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can	
QY	2050	CCACACAGCTCTCTCGCTAATCAGTTGATCAGCTTCCAACTCTGCAATTTGCTCTT	2109	CC	be used for preventing, treating or ameliorating a disorder related to	
DB	341	ProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu	360	CC	Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,	
QY	2110	TCTTCTTTGTTAGCAACATGAGTTACCAATCTCTTACCAATCTCTCAGCGCAGAT	2169	CC	lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,	
DB	361	SerSerLeuLeuAlaAsnAlaValThrAsnProProThrAsnProProAlaGlnAsp	380	CC	pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical	
QY	2170	TCTCATCTCGAGTCATTTGTAGCACAACCTGCTGTTCTGTTCAATTTAGTGGCCTATC	2229	CC	cancer, infertility, arteriosclerosis and atherosclerosis. The products	
DB	381	SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle	400	CC	can also be used for detection and diagnosis. Sequences AAV16740-Y16752	
QY	2230	TTTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTAATCAA	2289	XX	represent Chlamydia HMW protein fragments	
DB	401	PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGln	420	QY	Sequence 325 AA;	
QY	2290	AAATCAATCTCGTAATTTACGTTAGGAGTAAAGCCCAAGCTAATGCCCATCAGAT	2349	QY	Alignment Scores:	
DB	421	LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAsp	440	QY	Pred. No.: 9.67e-146 Length: 325	
QY	2350	TTGACTCTAGGGAATGAGTCCCTAAGTATGGCTATCAAGGAAGCTTGAAGCTT	2403	QY	Score: 1735.00 Matches: 325	
DB	441	LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeu	458	QY	Percent Similarity: 100.0% Conservative: 0	
				QY	Best Local Similarity: 100.0% Mismatches: 0	
				QY	Query Match: 22.0% Indels: 0	
				QY	DB: 0 Gaps: 0	
				QY	US-10-701-844-1 (1-4435) x AAV16752 (1-325)	
				QY	2443 CTGAAAGCTACATGGAGTAAACTGGGTATTAATCTCTGGGCTGAGCGAGTAGCTTCTTTG	2502
				DB	1 LeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu	20
				QY	2503 GTTCCAAATAGTTTATGGGATCCATTTAGATATAGATCTCGGCATTCAGCAATTCAA	2562
				DB	21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln	40
				QY	2563 GCAAGTGTGAGCGGCTCTTATTCGAGGATTAATGGCTTCTGAGTTTCGAATTC	2622
				DB	41 AlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhe	60
				QY	2623 TTCTATCATCACCGCGATGCTTTAGGTCAGGATATCGGTATATAGTGGGGTATTTC	2682

```
Db 61 PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSer 80
QY 2683 TTAGAGCAAACTCTACTTTGGATCATCGATGTTTGGTCTAGCAATTTACCGAAGTATTT 2742
Db 81 LeuGlyAlaAsnSerTyrPheGlySerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAGATATGATGTGTGCTTCCAAATCATCATGCTTGCATGATGATCCGTT 2802
Db 101 GlyArgSerLysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCCACAAAGCTTATGTGATCTATTTGTCGGAGATCGGTTATTCGCT 2862
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg 140
QY 2863 GCTACTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAGAGC 2922
Db 141 AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer 160
QY 2923 GATGTCGTTGGGATAATAACTGTCCTGCTGGTGAGAGATTGGAGCGGATTTACCGATTGTG 2982
Db 161 AspValArgTrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180
QY 2983 ATTACTCCCATCTAAGCTCTATTGATGATGATGCTGCTTCTGTCGCAAGCTGAGTTTCT 3042
Db 181 IleThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATGCCGATCATGAATCTTTTACAGAGGAAGCGGATCAAGCTCGGGCATTTCAAGAGCGGA 3102
Db 201 TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly 220
QY 3103 CATCTCTAAATCTATCATGTTCTGTTGGAGTGAAGTTTGATGATGATGTTCTAGTACACAT 3162
Db 221 HisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerThrHis 240
QY 3163 CCTAATAATATAGCTTTATGCGGCTATATCTGATGCTTATCGCACCATCTCTGCT 3222
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260
QY 3223 ACTGAGACAACGCTCTCTCCATCCATCAAGACATGACACAGATGCGCTTTCATTAGCA 3282
Db 261 ThrGluThrThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAla 280
QY 3283 AGACATGGAGTTGTGGTTAGAGATCTATGATGCTTCTCTAAAGTAATATAGAATA 3342
Db 281 ArgHisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 300
QY 3343 TATGCCATGGAAGATATGATGATGATGATGCTTCTCGAGGCTATGTTTGTGATGTCAGGA 3402
Db 301 TyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGly 320
QY 3403 AGTAGATCCGGTTC 3417
Db 321 SerArgValArgPhe 325
```

RESULT 35

```
AA37234
ID AAY37234 standard; protein; 708 AA.
```

AC AAY37234;

XX 07-OCT-1999 (first entry)

DE Chlamydia trachomatis cellular envelope protein.

DE Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW bartholinitis; pneumonia; venereal lymphogranulomatosis.
XX Chlamydia trachomatis.
XX WO928475-A2.

```
XX 10-JUN-1999.
XX 27-NOV-1998; 98WO-IB001939.
XX 28-NOV-1997; 97FR-00015041.
XX 17-DEC-1997; 97FR-00016034.
XX 04-NOV-1998; 98US-0107077P.
XX (GEST ) GENSET.
XX Griffais R;
XX WPI; 1999-371125/31.
XX Genome sequence of Chlamydia trachomatis.
XX Disclosure; Page 993-994; 1755pp; English.
XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma, such as
CC paratrachoma, and inclusion conjunctivitis; genital diseases, such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihhepatitis, bartholinitis; pneumonia in breast feeding infants; and
CC venereal lymphogranulomatosis. The polypeptides of the invention may be
CC of use in treating these diseases
XX SQ Sequence 708 AA;
Alignment Scores:
Pred. No.: 9,19e-144 Length: 708
Score: 1715.00 Matches: 331
Percent Similarity: 99.7% Conservative: 0
Best local Similarity: 99.7% Mismatches: 0
Query Match: 21.8% Indels: 1
DB: 2 Gaps: 0
US-10-701-844-1 (1-4435) x AAY37234 (1-708)
QY 3442 GTGTTAGCGATGCGCTTTTCTTGTGATCATCATCATTTTGTGTTTGTGTTGCT 3501
Db 1 ValLeuAlaMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 20
QY 3502 TCCTATTCTGATGCTTCCGAGCTCTCTCAAGTGTGTAAGCCCTTAATGTAACTCTCT 3561
Db 21 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 40
QY 3562 TTTAAGGAGAGCATGTTTACTTGAATGAGAGATCGCTGCTTTTGTCAATGTCTATGAGGA 3621
Db 41 PheLysGlyAspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 60
QY 3622 GCTGAAGAAGGTTTCCATTTATCTCAGCTAATGCGGCAATTTTAACGATACCGGCAAAAC 3681
Db 61 AlaGluGluGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 80
QY 3682 CATACATTATCATTTTACAGATTCTCAAGGCGCAGTTCTTCAAAATATATGCTTTCATTTCA 3741
Db 81 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 100
QY 3742 GCAGGAGAGACATTTACTCTGAGAGATTTTTCGAGTCTGATGTTTTCGAAAAATGTTTCT 3801
Db 101 AlaGlyGluThrLeuThrLeuArgAspPheSerSerLeuMetPheSerLysAsnValSer 120
QY 3802 TCCGGAAGAAAGGAAATGATCTCCCGGAAAACCGTACGATTTTTCGGGACGACGAGTG 3861
Db 121 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 140
QY 3862 ATTTTCTGGGATAACTCCGTGGGTATTTCTCTCTTATCTACTGTGTCACCTCATCATCA 3921
```

Db 141 llePheTrpAspAsnSerValGlyTyrSerProLeuSerThrValProThrSerSer 160
 QY 3922 ACTCCGCTGCTCCAAACAGTTAGTGTCCGAAAGGCTCTATTTTCTAGAGACT 3981
 Db 161 ThrProAlaProThrValSerAspAlaArgLysGlySerIlePheSerValGluThr 180
 QY 3982 AGTTTGGAGATCTCAGGCTCAAAAAGGGGTCATGTTTCGATAAATCCCGGGAATTC 4041
 Db 181 SerLeuGluIleSerGlyValLysGlyValMetPheAspAsnAsnAlaGlyAsnPhe 200
 QY 4042 GAAACAGATTTTTCGAGGTAAAGATAATAATATGCTGTGGTGGAGGAGTGGTTCGCC 4101
 Db 201 GlyThrValPheArgGlyLysAsnAsnAlaGlyGlyGlySerGlySerAl 220
 QY 4102 TACACCATCAAGTACGACTTTTACAGTTAAACTGTAAGGAAAGTTTCTTTCACAGA 4161
 Db 220 aThrProSerSerThrThrPheThrValLysAsnCysLysGlyValSerPheThrAs 240
 QY 4162 TAACGTAGCTCTTCGGAGGCGGAGTGTGTTTATAAGGCATTTGCTTTTCAAGACAA 4221
 Db 240 pAsnValAlaSerCysGlyGlyValValLysGlyValLysGlyValLysPheLysAspAs 260
 QY 4222 TGAAGGAGGCATATTTCTCCGAGGGAACACAGCATACGATGATTTAAGGATTTCTGCTGC 4281
 Db 260 nGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspLeuArgIleLeuAlaAl 280
 QY 4282 TACTAATCAGGATCAGAAATACGAGACAGAGCGGTGGAGGATTTTGTCTCTCCAGA 4341
 Db 280 aThrAsnGlnAspGlnAsnThrGluThrGlyGlyGlyGlyValLysCysSerProAs 300
 QY 4342 TGAATCTGAAGTTTGAAGGCAATAAAGTTCTATTGTTTGTATTACAACTTTTCAAA 4401
 Db 300 pAspSerValLysPheGluGlyAsnLysGlySerIleValPheAspTyrAsnPheAlaL 320
 QY 4402 AGCAGAGCGGGAAGCATCTCAACGAAAGAAATTC 4435
 Db 320 sGlyArgGlyGlySerIleLeuThrLysGluPhe 331
 RESULT 36
 AAG83274
 ID AAG83274 standard; protein; 631 AA.
 XX AC AAG83274;
 DT 05-SEP-2001 (first entry)
 XX DE Chlamydia trachomatis PmpH(N-term) fusion protein.
 XX KW Chlamydia; vaccine; infection; fusion protein; antigen;
 KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
 KW acute respiratory tract infection; Capi; CT529; OMCB;
 KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
 XX OS Chlamydia trachomatis.
 XX PN W0200140474-A2.
 XX PD 07-JUN-2001.
 XX PF 04-DEC-2000; 2000WO-US032919.
 XX PR 03-DEC-1999; 99US-00454684.
 XX PR 19-APR-2000; 2000US-00556877.
 XX PR 20-JUN-2000; 2000US-00598419.
 XX PA (CORI-) CORIXA CORP.
 XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
 XX WPI; 2001-374831/39.
 XX PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
 PT inflammatory disease, trachoma, acute respiratory tract infections,

PT atherosclerosis and heart disease.
 XX Claim 70; Page 274-276; 295pp; English.
 CC The present sequence is provided in a specification relating to compounds
 CC and methods for the treatment and diagnosis of chlamydial infection. The
 CC compounds provided include polypeptides and fusion proteins comprising
 CC immunogenic portions of Chlamydia antigens and DNA sequences encoding
 CC such polypeptides. They are useful for vaccinating against chlamydial
 CC infection, which causes pelvic inflammatory disease, trachoma, acute
 CC respiratory tract infections, atherosclerosis and heart disease
 XX Sequence 631 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1,21e-132 Length: 631
 Score: 1590.50 Matches: 316
 Percent Similarity: 88.7% Conservative: 7
 Best Local Similarity: 86.8% Mismatches: 19
 Query Match: 20.2% Indels: 22
 DB: 4 Gaps: 4
 US-10-701-844-1 (1-4435) x AAG83274 (1-631)
 QY 3391 TTGAGTGCAGGAAGTAGAGTCCGGTTCTAAATAATAT-----TGGTTAGATAGTTAA 3441
 Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTrpArgPro----- 146
 QY 3442 GTGTAGCAATGCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTAGCTTGTGTTGTGT 3501
 Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164
 QY 3502 TCTATTTCGTATGGAATCCGAGCTCTCTCAAGTGTAAACGGCTAATGTAACCACTCTCT 3561
 Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 184
 QY 3562 TTTAAGGAGACATGTTTACTTGAATGGAGACTCGCTTTTGTCTCAATGCTATGAGGA 3621
 Db 185 PheLysGlyAspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 204
 QY 3622 GCTGAAGAAGGTTCTGATATCTCAGCTAATGGGACAAATTAACGATTACCGACAAAC 3681
 Db 205 AlaGluAsnGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224
 QY 3682 CATACATTATCATTTACAGATTCTCAAGGCCAGTCTTCAAAATTTATGCTTTCATTTC 3741
 Db 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 244
 QY 3742 GCAGGAGACACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCT 3801
 Db 245 AlaGlyGluThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSer 264
 QY 3802 TCGGGAAGAAAGGAATGATCTCCGGGAAACCGTGAGTATTTCCGAGCAGGCGGAGTG 3861
 Db 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284
 QY 3862 ATTTTCTGGGTAACCTCCGTTGGGTATCTCTTTTATCTACTGTCGCAACTCATCATCA 3921
 Db 285 IlePheTrpAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304
 QY 3922 ACTCCGCTGCT-----CCACAGCTTAGT 3945
 Db 305 ThrProProAlaProAlaProAlaProAlaAlaSerSerSerLeuSerProThrValSer 324
 QY 3946 GATGCTCGGAAACGGTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAA 4005
 Db 325 AspAlaArgLysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLys 344
 QY 4006 AAAGGGTCTATGTTCCGATATATTAATGCGGAAATTCGGAACAGTTTTTCCGAGGTGAAGAT 4065
 Db 345 LysGlyValMetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer 364
 QY 4066 AATAATANTGCTGTGTGGAGGAGTGGGTTCGGTTCACCATCAAGTACGACTTTTAC 4125

XX Chlamydial infection; Chlamydia; antiinfective; antiinflammatory;
KW antinfertility; cardiant; antiarteriosclerotic; ophthalmological;
KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
KW tubal obstruction; infertility; male infertility; ocular infection;
KW blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease.
XX
OS Chlamydia trachomatis.
XX
PN WO2003041560-A2.
XX
PD 22-MAY-2003.
XX
PF 05-NOV-2002; 2002WO-US035624.
XX
PR 06-NOV-2001; 2001US-00012256.
PR 05-DEC-2001; 2001US-00007693.
PR 15-JUL-2002; 2002US-00197220.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bhatia A, Gudezian J, Skeiky YAW, Maisonneuve JL, Barth B;
PI Probst P;
DR WPI; 2003-441771/41.
DR N-PSDB; ADD42667.
XX
PT New DNA and proteins comprising a portion of a Chlamydia antigen, useful
PT for diagnosing or treating Chlamydial infections, particularly as
PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
PT inflammatory disease.
XX
PS Disclosure; SEQ ID NO 95; 275pp; English.
XX
CC The present invention describes compounds and methods for diagnosing and
CC treating Chlamydial infection. Chlamydia polynucleotide and protein
CC sequences have antibiotic, antiinflammatory, antifertility, cardiant,
CC antiarteriosclerotic and ophthalmological activities, and can be used in
CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
CC compositions or methods from the present invention can be used for the
CC serodiagnosis or treatment of Chlamydial infections, particularly in
CC humans. The polynucleotides, proteins or compositions are particularly
CC useful for stimulating an immune response in a patient, or for
CC stimulating and/or expanding T cells specific for a Chlamydia protein.
CC Specifically, the polynucleotides, proteins or compositions are useful as
CC vaccines for treating or preventing Chlamydial infections including
CC pelvic inflammatory disease (which results in tubal obstruction and
CC infertility in women), male infertility, ocular infection (which may
CC cause blindness), acute respiratory tract infections, atherosclerosis, or
CC coronary heart disease. The present sequence is used in the
XX exemplification of the present invention.

SQ Sequence 1016 AA;

Alignment Scores:
Pred. No.: 2,15e-132 Length: 1016
Score: 1589.00 Matches: 311
Percent Similarity: 92.7% Conservative: 5
Best Local Similarity: 91.2% Mismatches: 10
Query Match: 20.2% Indels: 15
DB: 7 Gaps: 2

US-10-701-844-1 (1-4435) x ADD42682 (1-1016)

QY 3451 ATGCCCTTTTTCCTTGAGATCATCATTTGTTTTTAGCTGTGTGTCTATTTCG 3510

Db 1 MetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysSerTy-Ser 20

QY 3511 TATGATTCGGAGCTCTCCACTAGTGTTAACGCCTATGTAACCACTCCTTTTAAGGA 3570

Db 21 TyrGlyPheAlaSerSerProGlnValLeuThrProGlnValThrThrProPheLysGly 40

3571	QY	GACGATGTTTACTTGAATCGGAGCTCCGCTTTTGTCTCAATGCTCTATGTCAGGAGCTGAAGAA	3633
41	DB	AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn	60
3631	QY	GGTTCGATTATCTCAGCTAATGGCGACAATTTAAACGATTACCGGACAAACCATACATTA	3690
61	DB	GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyIleAsnHisThrLeu	80
3691	QY	TCATTTACAGATTCTCAAGGCCAGTCTTCAAAATATATGCCCTTCATTTACAGCAGAGAG	3750
81	DB	SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu	100
3751	QY	ACACTTACTCTGAGAGATTTTTCGAGTCTGTAGTCTTCGAAATAATGTTCTTCGCGAGAA	3810
101	DB	ThrLeuThrLeuIysAspPheSerSerLeuMetPheSerIysAsnValSerCysGlyGlu	120
3811	QY	AAGGGAATGATCTCCGGGAAAACCGTAGTATTTCGGGACGAGCGAAGTGATTTTCTCG	3870
121	DB	LysGlyMetIleSerGlyIysThrValSerIleSerGlyAlaGlyGluValIlePheTrp	140
3871	QY	GATAACTCCGTGGGGTATCTCTTTATCTTACTGTGCGCAACCTCATCATCACTCCGCCT	3930
141	DB	AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro	160
3931	QY	GCT-----CCACAGTTACTGTAGTCTCGG	3954
161	DB	AlaProAlaProAlaProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArg	180
3955	QY	AAAGGCTATTTTTCGTAGACACTAGTTCGAGATCTCAGGCGCTCAAAAAGGGGTC	4014
181	DB	LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValIysGlyVal	200
4015	QY	ATGTTCCGATAATATGCGCGGAATTCGGAACAGTTCGAGGTAAAGAATAATAATAAT	4074
201	DB	MetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsn	220
4075	QY	GCTGGTCGTGGAGCGAGTGGTCCGCTCACCATCAAGTACGACTTTTACAGTTAAAAA	4134
221	DB	AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValIysAsn	238
4135	QY	CTGTAAAGGGAGAACTTCCTTTTCAGACATACTAGTCCCTTCGCGAGCGGAGTGGTTTA	4194
238	DB	nCysLysGlyIysValSerPheThrAspAsnValAlaSerCysGlyGlyGlyValValTy	258
4195	QY	TAAAGGCATTGTCTTTTCAAAGACAAATGAAGAGGAGCATATTCCTCGAGGGAAACAGC	4254
258	DB	rLysGlyThrValLeuPheIysAspAsnGluGlyIlePhePheArgGlyAsnThrAl	278
4255	QY	ATACAGATGATTAGGATTCCTGCTGCTACTTAATCAGGATCAGAAATACGAGACAGGAGG	4314
278	DB	atyAspAspLeuGlyIleLeuAlaThrSerArgAspGlnAsnThrGluThrGlyGly	298
4315	QY	CGGTGGAGGAGTATTTTGCTCTCCAGATGATTCTGTAAAGTTTCAAGCCCAATAAGGTTTC	4374
298	DB	yGlyGlyValIleCysSerProAspAspSerValIysPheGluGlyAsnLysGlyLys	318
4375	QY	TATTGTTTTGATTACAACCTTTGCAAAAGGCAGAGCGGAGCATCTTACGAAAGAATT	4434
318	DB	rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPhe	338
4435	QY	C 4435	
338	DB	e 338	

Search completed: May 13, 2006, 09:59:52
Job time : 938 secs

THIS PAGE BLANK (USPTO)

Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 11 Gaps: 1

US-10-701-844-1 (1-4435) x US-11-103-957-9 (1-1013)

Qy 382 ATGCAAGCTCTTCCCAAGTCTCTTCTTCAATGATTTAGCTTATCTTGTGCTCT 441
Db 1 MetGlnThrSerPheHisLeuPheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
Qy 442 TTAATGGGGGGGATATGACGAGAAATCATGTTCTCTCAAGGAATTCAGATGGGAG 501
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
Qy 502 ACGTTAACTGATCATTTCCCTATCTGTATAGGATCCGAGTGGGACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTGAGGAGAGTTAAACATTAAAAATCTTGACAAATCTATTGCAAGCTTTGCTTTAAGT 621
Db 61 SerIlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
Qy 622 TGTCTGGGAACCTATTAGGAGTTTACTGTTTGGGAGGAGACACTGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACCGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTACTATTAGGGTTTTAAAGAAATATCTTTTCCAAATTCATTAATCTTACCTGCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTGCTGCTGCAACGACTAATAAGGTTAGCGAGCTCCGAGTCCGACGACATCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
Qy 862 AATGCTACTATTATTCTTAAACAGACTCTTTGTTTACTCAATAATGAGAAGTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuLeuLeuAsnGlnLysPheSerPhe 180
Qy 922 TATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
Qy 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTACCAGTTCTCTGCTATGCTAACGAGGCTCTATTGCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy 1102 GTTGAGGAGTAAGAGGGGAGGATTCGTCTGTTCAGGATGGGAGGAGGAGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaValGlnAspGlyGlnGlnGlyValSer 260
Qy 1162 TCATCTACTTCAAGAGATCCAGTAGTAGTTTTCAGAAATATCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Qy 1222 GATGGGACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAAGCTTCTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Qy 1282 AATAATGGAACACCTGTTTCTCAACAAATGTGCTTCTCTGTTTACATTTGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
Qy 1342 CAACCAACAGTGGGACGGCTTCTAATAGGAGTAATAATTACGGAGATGGAGGACTATC 1401

Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
Qy 1402 TTCTGTAAAGATGTGTGCGCAA--GCAGGATCCCAATAACTCTCGATCAGTTCTCTTGTAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
Qy 1459 GGAGAGGAGTAGTTTCTTTTAGTACAAATGCTGCTGGGAAAGGGGAGCTATTATAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
Qy 1519 GCCAAAAGCTCTCGTGTCTTAACGTGGCCCTGACAAATTTTAAAGAAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
Qy 1579 GATCGTGAGCGATTTATTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATTTATCGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGluLeuSerLeuSerAlaAspTyrGly 420
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAAATGCTGCGATGTTAAT 1698
Db 421 AspileIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
Qy 1699 GGCGTAACTGTCTCACAGCCATTTGATGGATCGGAGGGGAGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
Qy 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
Qy 1819 AACGAGCAGCGAGTCTTCCAACTCTTAAATAATTAAACGATGGTGAAGGATACACAGGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
Qy 1879 GATATTGTTTGTCTTAAAGCAGACTTTGTACCAAAATGTACCAAAATGTACGATAGACAAAGA 1938
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Qy 1939 AGGATTGTTCTCTGAAAAGGCCAAATTAATCAGTGAATCTCTAAAGTCAGACAGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Qy 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACACCAACCAACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
Qy 2059 CCTCTCGCTAATCAGTTGATCAGCTTCCCAATCTGCAATTTGCTCTCTTCTCTTTG 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
Qy 2119 TTAGCAACAATGAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTTCTCATCCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
Qy 2179 CGAGTCAATGTGTAGCACAACTGCTGTTCTGTTTACAATTAGTGGGCCCTATCTTTTTCAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy 2239 GATTTGGATGATACAGCTTATGATAGGTATCATTTGGCTAGGTCTTAATCAAAAATATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspIleLeuGlySerAsnGlnLysIleAsp 640
Qy 2299 GTCTCGAAATACAGTTAGGAGACTAAGCCCCCAGCTAATGCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Qy 2359 GGGAAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAAGCTTGGCTGGATCTTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTrpAspProAsn 680
Qy 2419 ACAGCAATAATGTTCTTATCTACTGAAAGCTACATGACTTAAACCTGGGTATATATCTCT 2478

```
Db      681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
Qy      2479 GGGCTGAGCGAGTAGCTTCTTTGTTCCAAATAGTTTATGGGATCCATTTTACATATA 2538
Db      701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
Qy      2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCCAGGATTA 2598
Db      721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy      2599 TGGGTTCTCTGGAGTTCCGAATTTCTTATCATGACCCGCGATGCTTTAGGTCCAGGATAT 2658
Db      741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy      2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTACTTGGATCCATCGATGTTT 2718
Db      761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerMetPhe 780
Qy      2719 GGTCTAGCATTTACCGAAGTATTGTGTAGATCTAAAGATTATGTAGTGTCTTCCCAAT 2778
Db      781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy      2779 CATCATGCTGTCATAGGATCCGTTTATCTATCTACCCAAAGCTTTATGTGCATCCTAT 2838
Db      801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy      2839 TTGTTCCGAGATGCGTTTATCCGTCTAGCTAGCTAGCGGTTTGGGAATCAGCATATGAACCC 2898
Db      821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy      2899 TCATATACATTTGCAGAGCAGAGCATGTTCTGGGATAATAACTGTCTGGCTGGAGAG 2958
Db      841 SerTyrThrPheAlaGluGluSerAspValAlaGrpAspAsnAsnCysLeuValGlyGlu 860
Qy      2959 ATTGAGCGCGGATTACCGATTGTGATTCTCTCATCTAAGCTCTTATTGAATGAGTTGCGT 3018
Db      861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy      3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGACCGAT 3078
Db      881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy      3079 CAAGCTCGGGCATTCAGAGCGGACATCTCCATAATCTATCAGTCTCTGTGGAGTGAAG 3138
Db      901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy      3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGGCGGCTTATATCTGT 3198
Db      921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy      3199 GATGCTTATCGCAACCATCTCTGTTACTGAGACAAACGCTCTCTATCCCATCAGAGCATGG 3258
Db      941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy      3259 ACACAGATGCTTCTTATAGCAACATGAGTGTCTGTTAGAGATCTATGATGCT 3318
Db      961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy      3319 TCTCTAACAGTAAATAGAGTATATGGCCATGGAAGATGATGAGTATCGAGATGCTTCT 3378
Db      981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy      3379 CGAGGCTATGGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db      1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

RESULT 2

```
US-10-701-844-1 (1-4435) x US-11-018-868-21 (1-1013)
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018.868
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-21
Alignment Scores:
Score: 0 5131.50 1013
Percent Similarity: 98.8% Conservatives: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 11 Gaps: 1
Qy      382 ATGCAAAAGCTCTTTCCATAAGTTCTTTCTTTCAATGATTCTAGCTTATTCTTCTGCTCT 441
Db      1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
Qy      442 TTAATATGGGGGGATATGCACAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
Db      21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
Qy      502 ACGTAACTGTATCATTTCCCTTACTGTATAGGAGATCCGAGTGGGACTCTGTTT 561
Db      41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy      562 TCTCGAGGAGAGTTAAACATTAAAAAATCTTGACAAATCTTATTCAGCTTTGCTTTAAGT 621
Db      61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
Qy      622 TGTTTTGGAACTTATTAGGAGTTTTACTGTTTTAGGAGAGGACACTCGTTGACTTTC 681
Db      81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy      682 GAGAACATGCACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACATG 741
Db      101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy      742 TTTACTATTGAGGTTTTAAAGAAATTTATCTTTTCCAAATTCCTTCAATTCATTACTTGCCTGA 801
Db      121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy      802 CTGCTGCTGCACACGACTAATAAGGGTAGCCAGACTCCGACGACAAACATCTACACCGTCT 861
Db      141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
Qy      862 AATGGTACTATTATTCTAAAACAGATCTTTGTTTACTCAATAATGAGAAGTTCTCATTC 921
Db      161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
Qy      922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTTATAGATGCTAAGAGCTTTAAACGGTTCAA 981
Db      181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
Qy      982 GGAATTTAGCAAGCTTTGTGTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db      201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy      1042 CAAGTAGTCACCAAGTTTCTCTGCTATGGCTTAACGAGGCTCTTATGCTTTGTAGCGAAT 1101
Db      221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
```

QY 1102 GTTCAGAGTAAGAGGGGAGGATTCCTGCTGTTCCAGGATGGCCAGCAGGAGTGTCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyLeAlaAlaValGlnAaspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGAGTCCAGTAGTAAAGTTTTCAGAAATACATCGCGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAaspProValValSerPheSerArgAasnThrAlaValGluPhe 280
QY 1222 GATGGAAAGTAGCCCGAGTAGGAGGAGGATTTACTCTCTCAGGAAAGCTTGCTTTCCTG 1281
Db 281 AspGlyAasnValAlaArgValGlyGlyGlyLeTySerTyGlyAasnValAlaPheLeu 300
QY 1282 AATATGAAAAACCTTGTTCTCAACAATGTTGCTTCTCTGTTTCAATTCGTGCTAAG 1341
Db 301 AasnAenGlyLeThrLeuPheLeuAasnValAlaSerProValTyLeAlaAlaGlu 320
QY 1342 CAACCAACAGTGGACAGCTTCTAATACAGAGTAATAATACGAGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAasnGlyAlaSerAasnThrSerAasnTyGlyAaspGlyGlyAlaLe 340
QY 1402 TTCTGTAAAGATGTGCGCAA---GCAGGATCCAAATACTCTGGATCAGTTTCTCTTCAT 1458
Db 341 PheCysLeAasnGlyAlaGlnAlaAlaGlySerAasnSerGlySerValSerPheAasp 360
QY 1459 GGAGGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGAAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAasnValAlaAlaGlyLeGlyGlyAlaLeTy 380
QY 1519 GCCAAAAAGCTCCGGTCTTAACGTGCGCTGTACAAATTTTAAAGAAATATCCCTAAT 1578
Db 381 AlaLeysLeuSerValAlaAasnCysGlyProValGlnPheLeuGlyAasnLeAlaAasn 400
QY 1579 GATGTGGAGCGCAATTTATTTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATATGGA 1638
Db 401 AspGlyGlyAlaLeTyLeuGlyGlyLeuSerGlyLeuLeuSerLeuAlaAaspTyGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTGCCGATGTTAAT 1698
Db 421 AspLeilePheAaspGlyAasnLeuLeysArgThrAlaLeysGluAasnAlaAlaAaspValAasn 440
QY 1699 GCGCTAACTGCTCTCAAGCCATTTTCGATGGGATCGGAGGAGGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaLeSerMetGlySerGlyGlyLeysLeThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGCGCAACGGAAT 1818
Db 461 ArgAlaLeysAlaGlyHisGlnLeuPheAasnAaspProileGluMetAlaAasnGlyAasn 480
QY 1819 AACAGCCAGCCAGTCTTCCAAACTTCTAAAAATTAACGATGGTGAAGGATACACAGG 1878
Db 481 AasnGlnProAlaGlnSerSerGluProLeuLeysLeAasnAaspGlyGlyTyThrGly 500
QY 1879 GATATTGTTTGTCTAAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGA 1938
Db 501 AspLeileValPheAlaAasnGlyAasnSerThrLeuTyGlnAasnValThrleGluGlnGly 520
QY 1939 AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
Db 521 ArgLeileValLeuArgGluLeysAlaLeysLeuSerValAasnSerLeuSerGlnThrGlyGly 540
QY 1999 ACTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACACACACACAG 2058
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAaspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTGCGCGCTTAATCAGTTGATCAGCTTTTCCATCTCATTGTTCTTCTCTCTTCTTG 2118
Db 561 ProProAlaAlaAasnGlnLeuLeThrLeuSerAasnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAAAACAATGCAATTCCTCTACCAATCTCCAGCGCAAGATTCATCTCT 2178
Db 581 LeuAlaAasnAasnAlaValThrAasnProProThrAasnProProAlaGlnAaspSerHisPro 600
QY 2179 GCAGTCATTTGGTAGCAAACTGCTGTTCTGTTCATTAATTTAGTGGGCTATCTTTTGTAG 2238

Db 601 AlaileileGlySerThrThrAlaGlySerValThrleSerGlyProilePhePheglu 620
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTAATCAAAAAATCAAT 2298
Db 621 AspLeuAaspThrAlaTyAspArgTyAspTrpLeuGlySerAasnGlnLeysleAasp 640
QY 2299 GTCTGAAATTTACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCCCATCAGATTGACTCTA 2358
Db 641 ValLeuLeysLeuGlnLeuGlyThrGlnProSerAlaAasnAlaProSerAaspLeuThrLeu 660
QY 2359 GGGAAATCAGATGCTTACGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGATCCTAAT 2418
Db 661 GlyAasnGluMetProLeysTyGlyTyGlnGlySerTrpLeysLeuAlaTrpAaspProAasn 680
QY 2419 ACAGCAATAATGTGCTTATCTCTGAAAGCTACATGAGCTATAAACTGGGTATATCT 2478
Db 681 ThrAlaAasnAasnGlyProTyThrLeuLeysAlaThrTrpThrLeysThrGlyTyAasnPro 700
QY 2479 GGGCTGAGCGAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAasnSerLeuTrpGlySerleLeuAaspIle 720
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTGTGCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaileGlnAlaSerValAaspGlyArgSerTyCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGGATTCGAATTTCTCTCATCATGACCCGATGCTTTTAGGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAasnPheTyHisAaspArgAaspAlaLeuGlyGlnGlyTy 760
QY 2659 CGGTATATTAGTGGGGTTTATCTTCTAGGACCAAACTCTACTTTGGATCATCGATGTT 2718
Db 761 ArgTyLeileSerGlyGlyTySerLeuGlyAlaAasnSerTyPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTTAAAGATATTAGTAGTGTGCTGTTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerTyAspTyValValCysArgSerAasn 800
QY 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTACCCCAACAGCTTTTATGAGATCTCTAT 2838
Db 801 HisHisAlaCysleileGlySerValTyLeuSerThrTyLeysGlnAlaLeuCysGlySerTy 820
QY 2839 TTGTTCCGAGATGCGTTTATCGTGCTAGCTACGGGTTTGGGATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAaspAlaPheileArgAlaSerTyGlyPheGlyAasnGlnHisMetTyThr 840
QY 2899 TCATATACATTTGACAGAGAGCGATGTTGTTGGGATAATAACTGCTGCTGCGAGAG 2958
Db 841 SerTyThrPheAlaGluGluSerAaspValArgTrpAaspAasnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTAACCGATTGTGATTACTCCATCTAAGCTCTATTGAATCAGTTGCGT 3018
Db 861 IleGlyValGlyLeuProileValleileThrProSerTyLeuTyLeuAasnGluLeuAsg 880
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGAAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyAlaAaspHisGluSerPheThrGluGluGlyAasp 900
QY 3079 CAAGCTCCGGCAATTCAGAGGCGGACATCTCTTAAATCTATCATGTTCTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheAargSerGlyHisleuMetAasnLeuSerValProValGlyValTy 920
QY 3139 TTTGATCAGATGTTCTAGTACACATCCTTAATAATATAGCTTTATGGCGCTTATCTCT 3198
Db 921 PheAaspArgCysSerSerThrHisProAasnTySerPheMetGlyAlaTyIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGCTACTGAGACACGCTCTCTATCCCATCAAGAGCATCG 3258
Db 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGlnThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGTGTGTTAGGATCTATCTATGTTGCT 3318

Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 Qy 3319 TCTCTAAACAAGTAATATAGAGTATATGCCATGGAAGATATGAGTATCGAGATCTTCT 3378
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 Qy 3379 CGAGGCTATGGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerIlysValArgPhe 1013

RESULT 3

US-11-103-957-41
 ; Sequence 41, Application US/1103957
 ; Publication No. US20050281847A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berthet, Francois-Xavier Jacques
 ; APPLICANT: Lobet, Yves
 ; APPLICANT: Poolman, Jan
 ; APPLICANT: Verlant, Vincent Georges Christian Louis
 ; TITLE OF INVENTION: Vaccine Composition
 ; FILE REFERENCE: B45261
 ; CURRENT APPLICATION NUMBER: US/11/103,957
 ; CURRENT FILING DATE: 2005-04-12
 ; PRIOR APPLICATION NUMBER: US/10/467,534
 ; PRIOR FILING DATE: 2004-02-03
 ; PRIOR APPLICATION NUMBER: PCT/EP02/01356
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: GB 0103169.9
 ; PRIOR FILING DATE: 2001-02-08
 ; NUMBER OF SEQ ID NOS: 108
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 41
 ; LENGTH: 1016
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis
 US-11-103-957-41

Alignment Scores:
 Pred. No.: 5,85e-130 Length: 1016
 Score: 1589.00 Matches: 311
 Percent Similarity: 92.7% Conservativity: 5
 Best Local Similarity: 91.2% Mismatches: 10
 Query Match: 20.2% Indels: 15
 DB: 11 Gaps: 2

US-10-701-844-1 (1-4435) x US-11-103-957-41 (1-1016)

Qy 3451 ATGCGCTTTTCTTGAGATCTACATCATTTTGTGTTTGTGTTCTTATTCG 3510
 Db 1 MetPropHeserLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer 20
 Qy 3511 TATGATTCGAGCTCTCTCAAGTGTTAAGCCCTAAATGTAAACCACTCTTTTAAAGGA 3570
 Db 21 TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrProPheIysGly 40
 Qy 3571 GACGATGTTTACTTGAATCGAGCTCGCTTTTGTCAATGCTCTATCGAGAGCTGAAGA 3630
 Db 41 AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn 60
 Qy 3631 GGTTCGATTTATCTCAGCTAATGGCGCAATTTAAACGATTACCGGACAAAAACCATACATTA 3690
 Db 61 GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu 80
 Qy 3691 TCATTTACAGATTCTCAAGGCGCAGTCTTCAAAATTTATGCTTCATTTTTCAGCAGAGAG 3750
 Db 81 SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu 100
 Qy 3751 ACATTACTCTGAGAGATTTTTCGACTCTGATGTTTCTCGAAAAATGTTTCTTCGCGAGAA 3810
 Db 101 ThrLeuThrLeuIysAspPheSerSerLeuMetPheSerIysAsnValSerCysGlyGlu 120
 Qy 3811 AAGGGAATCATCTCCGGGAAAAACCGTGAGTATTTCCGGAGCAGCGCAAGTGATTTTCTCG 3870

Db 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrp 140
 Qy 3871 GATAACTCCGGGGGTATTTCTCTTTATCTACTGTGCCCAACTCATCATCAACTCCGCT 3930
 Db 141 AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro 160
 Qy 3931 GCT-----CCAACAGTTAGTATGCTCGG 3954
 Db 161 AlaProAlaProAlaProAlaAlaSerSerLeuSerProThrValSerAspAlaArg 180
 Qy 3955 AAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGACTCTCAGCGCTCAAAAAAGGGTC 4014
 Db 181 LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValIlysGlyVal 200
 Qy 4015 ATGTTTCGATAATAATGCGGGAATTTCCGAACAGATTTTTCGAGGTAAGAATAATAATAAT 4074
 Db 201 MetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsn 220
 Qy 4075 GCTGCTGCTGGAGCGAGTGGTTCGCTTACCATCAAGTACGACTTTTACAGTTAAAAA 4134
 Db 221 AlaGlySerGlySer-GlySerAlaThrThrProSer-----PheThrValIysAs 238
 Qy 4135 CTGTAAAGGGAAGTTCTTTTTCACAGATAACGTAGCTCTTCCGAGGCGGAGTGGTTTA 4194
 Db 238 nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyValValTy 258
 Qy 4195 TAAAGGCATTGTGCTTTTCAAAGCAATGAAGGAGGCATATTCTTCCGAGGGAACACAGC 4254
 Db 258 rLysGlyThrValLeuPheLysAspAsnGluGlyIlePhePheArgGlyAsnThrAl 278
 Qy 4255 ATACGATGATTTAAGGATTCTTGCTTACTAATCAGGATCAGATAATCGAGACAGGAGG 4314
 Db 278 aTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGl 298
 Qy 4315 CGTCGAGAGTATTTCCTCTCCAGATGATTCCTAAAGTTTGAAGGCAATAAAGGTTTC 4374
 Db 298 yGlyGlyValIleCysSerProAspAspSerValIysPheGluGlyAsnLysGlySe 318
 Qy 4375 TATTCGTTTTCATTACAACTTTTGAAGGAGGCGGAGCGGCAATCAAGAAAT 4434
 Db 318 rIleValPheAspTyrAsnPheAlaLysGlyArgGlySerIleLeuThrIysGluPh 338
 Qy 4435 C 4435
 Db 338 e 338

RESULT 4

US-11-018-868-22
 ; Sequence 22, Application US/11018868
 ; Publication No. US20060034871A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chiron Corporation
 ; APPLICANT: Grandi, Guido
 ; APPLICANT: Ratti, Giulio
 ; APPLICANT: Bonci, Alessandro
 ; APPLICANT: Finco, Oretta
 ; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
 ; FILE REFERENCE: 002441.00099 (PP23152.001)
 ; CURRENT APPLICATION NUMBER: US/11/018,868
 ; CURRENT FILING DATE: 2004-12-22
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 22
 ; LENGTH: 1016
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis
 US-11-018-868-22

Alignment Scores:
 Pred. No.: 5,85e-130 Length: 1016
 Score: 1589.00 Matches: 311
 Percent Similarity: 92.7% Conservativity: 5
 Best Local Similarity: 91.2% Mismatches: 10


```

Db      750 LysLeuGlnGlyPheSerGlyThrGlnAspGlyPheGluGluSerSerGlyGluLeuArg 769
Qy      308 GCATTCAAGACGGGACATCTCTAAATCTATCAGTCTCTGTGGAGTGAAGTTGATCGA 3147
Db      770 SerPheSerAlaSerSerPheArgAsnIleSerLeuProMetGlyIleThrPheGluLys 789
Qy      3148 TGTTCTAGTACACATCCTCAATAAATATAGCTTATGGGGCTTATCTGTGATCGCTTAT 3207
Db      790 LysSerGlnLysThrArgAsnTyrTyrPheLeuGlyAlaTyrIleGlnAspLeuLys 809
Qy      3208 GCGACCATCTCTGTTACTGAGACACGCTCTATCCCATCAGACACATGACACACAGAT 3267
Db      810 ArgAspValGluSerGlyProValValLeuLeuLysAsnAlaValSerTyrAspAlaPro 829
Qy      3268 GCCTTTCATTATGCAAGACATGGAGTTGTGTTAGAGGATCTATGATCTCTCTAACA 3327
Db      830 MetAlaAsnLeuAspSerArgAlaTyrMetPheArgLeuThrAsnGlnArgAlaLeu-- 848
Qy      3328 AGTAATATAGAGTATATGGCATGGACATGGAAGATATGATATCGAGATGCTCTCGAGGCTAT 3387
Db      849 HisArgLeuGlnThrLeuLeuAsnValSerTyrValLeuArgGlyGlnSerHisSerTyr 868
Qy      3388 GGTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db      869 SerLeuAspLeuGlyThrThrTyrArgPhe 878

RESULT 6
US-11-018-868-23
; Sequence 23, Application US/11018868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-23

Alignment Scores:
Pred. No.: 1,26e-58 Length: 878
Score: 771.50 Matches: 268
Percent Similarity: 40.9% Conservative: 145
Best Local Similarity: 26.5% Mismatches: 390
Query Match: 9.8% Indels: 207
Db: 11 Gaps: 35

US-10-701-844-1 (1-4435) x US-11-018-868-23 (1-878)
Qy      538 GATCCGAGTGGGACTACTGTTTTCTCGCAGGAGAGTTAAACATTAAATAATCTTCACAAT 597
Db      26 AspProLeuGlyGluThrAla-----LeuLeuThrIleAsnProAsnHis 40
Qy      598 TCTATTGCAGCTTTCCTTAAAGTTGTTTGGGAACCTTATTA----- 639
Db      41 ValValCysThrPhePheGluAspCysThrMetGluSerLeuPheProAlaLeuCysAla 60
Qy      640 -----GGGAGTTTACTGTTTATGGAGAGGACACTGTTGATCTTCGAG 684
Db      61 HisAlaSerGlnAspAspProLeuTyrValLeuGlyAsnSerTyrCysTrpPheValSer 80
Qy      685 AACATACGGACTCTTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTT 744
Db      81 LysLeuHisIleThrAspProLysGluAlaLeuPheLysGluLysGlyAsp-----Leu 98

```

```

Qy      745 ACTATTGAGGGTTTAAAGAATTATCTTTTCCAAATGCAATTCATTACTTCCCGTACTG 804
Db      99 SerIleGlnAsnPheArgPheLeuSerPheThrAspCysSerSer----- 113
Qy      805 CCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACAAACATCTACACCGCTAAT 864
Db      114 -----LysGluSerSerProSer----- 119
Qy      865 GGTACTATTATTCTAAACA---GATCTTTTGTACTCAATAATGAGAAGTTCTCATTC 921
Db      120 ---IleIleHisGlnLysAsnGlyGlnLeuSerLeuArgAsnAsnGlySerMetSerPhe 138
Qy      922 TATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db      139 CysArgAsnHisAlaGluGlySerGlyIleAlaIleSerAlaAspAlaPheSerLeuGln 158
Qy      982 GGAATTAGCAAGCTTTGTGCTCTCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db      159 HisAsnTyrLeuPheThrAlaPheGluGlnAsnSerSerLysGlyAsnGlyIleAlaIle 178
Qy      1042 CAAGTAGTCCACAGTTTCTCTGCTATGGCTTAACGAGGCTCTTATTCCTTTGTAGCGAAT 1101
Db      179 GlnAlaGlnThr---PheSerLeuSerArgAsnValSerProIleSerPheAlaArgAsn 197
Qy      1102 GTTCGAGAGTAAGAGGGGGAGGATGCTGCTGTTTCAGGATGGGCACGACGAGGTGCA 1161
Db      198 ArgAlaAspLeuAsnGlyIleAlaIleCysCys----- 208
Qy      1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCACAGAAATACTCGCGTAGAGTTT 1221
Db      208 ----- 208
Qy      1222 GATGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTTCTCTG 1281
Db      209 -----SerAsnLeuIleCysSerGlyAsnVal----- 217
Qy      1282 AATAATGAAAAACCTTGTCTTCTCAACAATGTTGCTTCTCTCTGTTTACATGCTGCTAAG 1341
Db      218 -----AsnProLeuPhePhe----- 222
Qy      1342 CAACCAACAAGTGGACAGGCTTCTAATACAGTAGTAATAATACGGAGATGGAGGACTATC 1401
Db      223 -----ThrGlyAsnSerAlaThrAsnGlyIleAlaIle 233
Qy      1402 TTCTGTAAGAATGGTGGCAGCAGGATCCAAATACTCTGGATCAGTTTCTTCTTGATGGA 1461
Db      234 CysCysIle-----SerAspLeuAsnThrSerGluLysGlySerLeuSerLeuAlaCys 251
Qy      1462 GAGGAGTAGTTTCTTTAGTAGCATGCTGCTGGGAAAGGGGAGCTATTATTTATGCC 1521
Db      252 AsnGlnGluThrLeuPheAlaSerAsnSerAlaLysGluLysGlyIleAlaIleTyrAla 271
Qy      1522 AAAAAGCTCTCGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGAT 1581
Db      272 LysHisMetValLeuArgTyrAsnGlyProValSerPheIleAsnAsnSerAlaLysIle 291
Qy      1582 GGTGAGCGCATTTATTTAGGAGAAATCTCGAGAGCTCAGTTTATCTCTGATTTATGAGAT 1641
Db      292 GlyGlyAlaIleAlaIleGlnSerGlySerLeuSerIleLeuAlaGlyGluGlySer 311
Qy      1642 ATTATTTTCGATCGGAATCTTAAAGAACAGCCAAAGAAATGCTGCCGATGTTTATGCG 1701
Db      312 ValLeuPheGlnAsnAsnSerGlnArgThrSerAspGlnGlyLeuValArg---AsnAla 330
Qy      1702 GTAACGTGCTCTCAACAGCCATTTTCGATGGATCGGAGGGGAAATAACGACATTAAAGA 1761
Db      331 IleTyrLeuGluLysAspAlaIle-----LeuSerSerLeuGlu 343
Qy      1762 GCTAAGACGGGCATCAGATTCTCTTTTAAATGATCCATC---GAGATGGCAACAAAGAAAT 1818
Db      344 AlaArgAsnGly---AspIleLeuPhePheAspProIleValGlnGluSerSerSerLys 362

```

1819 ACCAGCCAGCCAGCTCTTCCAACTTCTTAATAATTAACGATGGTGAAGGATACACAGG 1878
 363 GluSerProLeuProSerSerLeuGlnAlaSerValThrSerProThrProAlaThrAla 382
 1879 -----GATATTGTTTTGCT----- 1893
 383 SerProLeuValIleGlnThrSerAlaAsnArgSerValIlePheSerSerGluArgLeu 402
 1894 -----AATGAACGAGTACTTTGTACCAAAATGTTAGGATA 1929
 403 SerGluGluGluLysThrProAspAsnLeuThrSerGlnLeuGlnProIleGluLeu 422
 1930 GAGCAAGGAGGATGTTCTTCGTGAAGGAGGAAATATCAGTGAATCTCTAAGTCAG 1989
 423 LysSerGlyArgLeuValLeuLysAspArgAlaValLeuSerAlaProSerLeuSerGln 442
 1990 ACAGGTGGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAA 2046
 443 AspProGlnAlaLeuLeuIleMetGluAlaGlyThrSer----- 455
 2047 CCACCAACAAGCCTCTCCGCTAATCAGTTGATCAGCGTTTCCAAATCTGCAATTTGTCT 2106
 456 -----LeuLysThrSerSerAspLeuLysLeuAla 465
 2107 CTTTCTTCTTTGTTAGCAACAATGCGATTACGAATCCTCTACCAATCTCCAGCGCAA 2166
 466 ThrLeuSerIleProLeuHisSerLeu----- 474
 2167 GATTCTCATCTGCACTGCTAGTACACAACTGCTGTTCTGTTACAAATTAGTGGGCTT 2226
 475 -----AspThrGluLysSerValThrIleHisAlaPro 485
 2227 -----ATCTTTTGTAGGATTG---GATGATACAGCTTATGATAGG 2265
 486 AsnLeuSerIleGlnLysIlePheLeuSerAsnSerGlyAspGluAsnPheThrGluAsn 505
 2266 TATGATTGGCTAGGTTCTAATCAAAAATCAATGCTCGAAATTTACAGTTAGGCACTAAG 2325
 506 ValGluLeuLeuSerLysGluGlnAsn---AsnIleProLeuLeuThrLeu----- 521
 2326 CCCCAGCTAAATGCCCATCAGATTTGACTCTA-----GGGAATGAGATGCCCTAAGTAT 2379
 522 -----SerLysGluGlnSerHisLeuHisLeuProAspGlyAsnLeuSerHisPhe 539
 2380 GGCTATCAAGAAAGCTGGAAGCTTGGTGGGATCTTAATACAGCAATAATGTCCTTAT 2439
 540 GlyTyrGlnGlyAspThrPheSerTrp-----LysAspSerAspGluGly---His 556
 2440 ACTCTGAAGCTACATGGACTAAACTGGGTATATCTCTGGGCTGAGCGAGTAGCTTCT 2499
 557 SerLeuIleAlaAsnTrpThrProLysAsnTrpValProHisProGluArgGlnSerThr 576
 2500 TTGGTTCCTCAATAGTTTATGGGATCCATTTTATAGATATAGCTCTGGCATTCCAGCAAT 2559
 577 LeuValAlaAsnTrpLeuTrpAsnThrTrpSerAspMetGlnAlaValGlnSerMetIle 596
 2560 CAAGCAAGTGGATGGCGCTCTTATTTGCGAGGATATGGGTTTCTGGAGTTTCAAT 2619
 597 AsnThrIleAlaHisGlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSerAsn 616
 2620 TTCTTCTAT---CATGACCGGATGCTTTAGT---CAGGATATCGGTATATAGTGGG 2673
 617 LeuPheTrpAlaHisAspSerSerGlyLysPheProIleAspAsnTrpHisArgSerLeu 636
 2674 GGTATTCTCTTAGGAGCAAACTCTAC---TTTGGATCATCGATGTTTGGCTAGCAATTT 2730
 637 GlyTyrLeuPheGlyIleSerThrHisSerLeuAspAspHisSerPheCysLeuAlaAla 656
 2731 ACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTCGTCCCAATCATGCTTGC 2790
 657 GlyGlnLeuLeuGlyLysSerSerAspSerPheIleThrSerThrGlu----- 672
 2791 ATAGGATCCGTTTATCTATCTATACCAACAAGCTTTATGTGGATCTTATTTGTTCCGGAGAT 2850

673 ---ThrThrSerTyrIleAlaThrValGlnAlaGlnLeuAlaThrProLeuMet---Lys 690
 2851 GCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATAT---ACA 2907
 691 IleSerAlaGlnAlaCysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSer 710
 2908 TTTGCAGAGGAGGAGCGATGTTCTGGTGGATAATAACTGTCTGGCTGGAGAGATTGGAGCG 2967
 711 PheSerLysGluGlyPheGlySerTrpHisSerValAlaValSerGlyGluValCysAla 730
 2968 GGATTACCGATTGTAATCTCATCTAAGCTCTATTTGAATGAGTTGGCTCTTCTCGTG 3027
 731 SerIleProIleValSerAsnGlySerGlyLeuPhe---SerSerPheSerIlePheSer 749
 3028 CAAGCTGAGTTTCTTATCCGATCATGAATCTTTTACAGAGCAAGGCGATCAAGCTCG 3087
 750 LysLeuGlnGlyPheSerGlyThrGlnAspGlyPheGluGluSerSerGlyGluLeuArg 769
 3088 GCATTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTGGAGTGAAGTTTGAATCGA 3147
 770 SerPheSerAlaSerSerPheArgAsnIleSerLeuProMetGlyIleThrPheGluLys 789
 3148 TGTCTAGTACATCTCTAATAATATAGCTTTTATGGCGCTTATATCTGTGATGCTTAT 3207
 790 LysSerGlnLysThrArgAsnTyrTyrPheLeuGlyAlaTyrIleGlnAspLeuLys 809
 3208 CCGACATCTCTGGTACTGAGACAGCTCTCTATCCCATCAAGAGACATGACACAGAT 3267
 810 ArgAspValGluSerGlyProValValLeuLeuLysAsnAlaValSerTrpAspAlaPro 829
 3268 GCCTTTCTATTAGCAAGACATCGGAGTTGTGGTGGAGGATCTATGTATGCTTCTTAACA 3327
 830 MetAlaAsnLeuAspSerArgAlaTyrMetPheArgLeuThrAsnGlnArgAlaLeu--- 848
 3328 AGTAATATAGAATATATGCGCATGGAAGATAGATAGATATGAGATCTCTCGAGGCTAT 3387
 849 HisArgLeuGlnThrLeuLeuAsnValSerTyrValLeuArgGlyGlnSerHisSerTyr 868
 3388 GGTTCAGTGCAGGAGTAGATCGGTTTC 3417
 869 SerLeuAspLeuGlyThrThrTyrArgPhe 878
 RESULT 7
 US-11-103-957-13
 ; Sequence 13, Application US/11103957
 ; Publication No. US20050281847A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berthet, Francois-Xavier Jacques
 ; APPLICANT: Lobet, Yves
 ; APPLICANT: Poolman, Jan
 ; APPLICANT: Verlant, Vincent Georges Christian Louis
 ; TITLE OF INVENTION: Vaccine Composition
 ; FILE REFERENCE: B45261
 ; CURRENT APPLICATION NUMBER: US/11/103,957
 ; PRIOR FILING DATE: 2005-04-12
 ; PRIOR APPLICATION NUMBER: US/10/467,534
 ; PRIOR FILING DATE: 2004-02-03
 ; PRIOR APPLICATION NUMBER: PCT/EP02/01356
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: GB 0103169.9
 ; PRIOR FILING DATE: 2001-02-08
 ; NUMBER OF SEQ ID NOS: 108
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 964
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis
 ; US-11-103-957-13
 Alignment Scores: 5.88e-48 Length: 964
 Pred. No.: 649.50 Matches: 260
 Score:

QY	2470	TATAATCTCGGCGCTGACGAGTACTCTTCTTGGTTCCAAATAGATTATGGGATCC---	2526	APPLICANT: Bonci, Alessandro
DB	651	TyrValProSerProLysHisArgSerProLeuIleAlaAsnThrLeuTyrGlyAsnMet	670	APPLICANT: Finco, Oretta
QY	2527	ATTTTAGATATACGATTCGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTAT	2586	FILE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
DB	671	LeuLeuAlaThrGluSerLeuLysAsnSerAlaGluLeuThrProSerGlyHisProPhe	690	FILE REFERENCE: 002441.00099 (PP23352.001)
QY	2587	TGTCGAGGATATAGGTTCTCGAGTTTCGAATTCCTTCATCATGACCGCATGCTTTA	2646	CURRENT APPLICATION NUMBER: US/11/018,868
DB	691	Tyr---GlyIleThrGlyGlyLeuGlyMetMetValTyrGlnAspProArgGluAsn	709	CURRENT FILING DATE: 2004-12-22
QY	2647	GGTCAGGATATCGGTATATTAGTGGGCTTATCTCTAGGA-----GCAACTCC	2697	NUMBER OF SEQ ID NOS: 186
DB	710	HisProGlyPheHisMetArgSerSerGlyTyrSerAlaGlyMetIleAlaGlyGlnThr	729	SOFTWARE: PatentIn version 3.2
QY	2698	TACTTTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGA---CTCAA	2754	SEQ ID NO 19
DB	730	HisThrPheSerLeuLysPheSerGlnThrTyrThrLysLeuAsnGluArgTyrAlaLys	749	LENGTH: 964
QY	2755	GATTATGTAGTGTGTCCTCCATCATCATGCTTCATAGATCCGTTTATCTATCTACC	2814	TYPE: PRT
DB	750	AsnAsnVal-----SerSerLysAsnTyrSerCysGlnGlyGluMetLeuPheSerLeu	767	ORGANISM: Chlamydia trachomatis
QY	2815	CAACAAGCT-----TTATGTGATCCTATTTTGTTCGGAGATCGGTTTATC	2859	US-11-018-868-19
DB	768	GlnGluGlyPheLeuLeuThrLysLeuValGlyLeuTyrSerTyrGlyAspHisAsnCys	787	Alignment Scores:
QY	2860	CGTGCTAGCTACGGGTTTCGGAATCAGCATATCAAAACCTCATATATTCGAGGAG	2919	Pred. No.: 5,88e-48
DB	788	HisHisPheTyrThrGlnGly---GluAsnLeuThrSerGlnGlyThrPheArgSerGln	806	Score: 649.50
QY	2920	AGCGATGTTCTGTGGGATAATAACTCTCTCGCTGGAGAGATTCGCGGATTTACGATT	2979	Percent Similarity: 41.1%
DB	807	Thr-----MetGlyGlyAlaValPhePheAspLeuProMet	818	Conservative: 260
QY	2980	GTGATTACTCCATCTAAGCTCTATTGATGATGAGTTCGTCCTTCGTCGACGCTAGTTT	3039	Best Local Similarity: 25.8%
DB	819	LysProPheGlySerThrHisIleLeuThr-----AlaProPheLeuGlyAlaLeuGly	836	Mismatches: 389
QY	3040	TCTTATGCGCATCATCAATCTTTTACAGAGAGAGCGCATCAAGCTCGGGCATTCAGAGC	3099	Query Match: 8.2%
DB	837	IleTyrSerSerLeuSerHisPheThrGluValGlyAlaTyrProArgSerPheSerThr	856	DB: 43
QY	3100	---GGACATCTCTAAATATCATGTTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGT	3156	US-10-701-844-1 (1-4435) x US-11-018-868-19 (1-964)
DB	857	LysThrProLeuIleAsnValLeuValProIleGlyValLysGlySerPheMetAsnAla	876	QY 631 AACTTATTAGGAGTTTACTGTTTTCAGGAGAGACACTCTGTTGACT-----TTC
QY	3157	ACACAT---CCTAATAATATAGCTTTATGCGCGCTTATATCTGATGCTTATCGCACC	3213	DB 39 SerLeuSerAsnLysIleSerLeuThrGlyAspThrHisAsnLeuThrAsnCysTyrLeu
DB	877	ThrHisArgProGlnAlaTyrThrValGluLeuAlaTyrGlnProValLeuTyrA-rgln	896	QY 682 GAGACATACGG-----ACTTCTACAANTGGGCGAGCTCTA
QY	3214	ATCTCTGGTACTGAGACAACGCTCCATCCATCAGAGACATGAGCAACAGATGCGCTT	3273	DB 59 AspAsnLeuArgTyrIleLeuAlaIleLeuGlnLysThrProAsnGluGlyAlaAlaVal
DB	897	GluProGlyIleAlaAlaGlnLeuLeuAlaSerLysGlyIleTyrPheGlySerGlySer	916	QY 718 ACTAATAGCGTCTGCTGAGTCTTACTATTATGAGGGTTTAAAGAAATATCTCTTCC
QY	3274	CATTTAGACAGATGAGGTTGTGGTTAGAGGATCTATGATGCTCT-----	3321	DB 79 -----ThrIleThrAspTyrLeuSerPhePheAspThrGlnLysGluGlyIleTyrPhe
DB	917	ProSerSerArgHisAlaMetSerTyrLysIleSerGlnGlnThrGlnProLeuSerTyr	936	QY 778 AATGCAATTCATTACTTCGCTACTGCTGCTGCAACGACTAATAAGGGTACCCAGACT
QY	3322	CTAACAGTAAATAGAGTATATGCG 3348		DB 97 AlaLysAsn-----LeuThrProGluSerGlyGlyAlaIleGlyTyrAlaSer
DB	937	LeuThrLeuHisPheGlnTyrHisGly 945		QY 838 CCGACGACAACTACACCGCTCTAATGCTATTTATTTATCTTAAACAGATCTTTTGTGA
RESULT 8				DB 113 ProAsnSerProThrValGluIleArgAspThrIle-----GlyProValIlePhe
US-11-018-868-19				QY 898 CTCNATAATGAG-----AAGTCTCATCTATAGTAATTTAGTCTCTGGAGAT-----
; Sequence 19, Application US/11018868				DB 130 GluAsnAsnThrCysCysArgLeuPheThrTyrPargAsnProTyrAlaAlaAspValVal
; Publication No. US20060034871A1				QY 946 -----GGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGT
; GENERAL INFORMATION:				DB 150 ArgGluGlyGlyAlaIleHisAlaGlnAsnLeuTyrIleAsnHisAsnHisAspValVal
; APPLICANT: Chiron Corporation				QY 1000 GTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCACTTTC
; APPLICANT: Grandi, Guido				DB 170 GlyPheMetLysAsnPheSerTyrValGlnGlyGlyAlaIleSerThrAlaAsnThrPhe
; APPLICANT: Ratti, Guilio				QY 1060 TCTGCTATGGCTTAACAGGCTCTATGCTTGTAGCGAATCTT-----
				DB 190 ValValSerGluAsnGlnSerCysPheLeuPheMetAspAsnIleCysIleGlnThrAsn
				QY 1105 ---GCAGGAGTAGAGGGGAGGAGGATTGCTGCTGTTTTCAGGATGGGACGAGGAGGTGCA
				DB 210 ThrAlaGly---LysGlyGlyAlaIleTyrAla-----GlyThrSer
				QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATATCTCGGTAGAGTTT
				DB 223 AsnSerPheGlu-----SerAsnAsnCysAspLeuPhePhe
				QY 1222 GATGGGAACTACCGCCGAGTAGGAGGAGGAGTTTACTCC-----
				DB 235 IleAsnAsnAlaCysCysAlaGlyGlyAlaIlePheSerProLeCysSerLeuThrGly
				QY 1261 ---TACGGGAACGTTGCTTTCTCTGAAATAATGGAAACCTTGTCTTCTCAACAATGTTGCT


```
QY 3274 CATTAGCAAGACATGGAGTCTGCTAGGAGATCTATGATGCTTCT----- 3321
Db 917 ProSerSerArgHisAlaMetSerTyrIleSerGlnThrGlnProLeuSerTrp 936

QY 3322 CTAACAAGTAATATAGAGTATATGCG 3348
Db 937 LeuThrLeuHisPheGlnTyrHisGly 945

RESULT 9
US-11-103-957-15
; Sequence 15, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-15

Alignment Scores:
Pred. No.: 2,52e-43 Length: 1531
Score: 597.50 Matches: 293
Percent Similarity: 36.7% Conservative: 153
Best Local Similarity: 24.1% Mismatches: 474
Query Match: 7.6% Indels: 295
DB: 11 Gaps: 54

US-10-701-844-1 (1-4435) x US-11-103-957-15 (1-1531)
QY 430 TCTTGCTGCTCTTTAAATGGGGGGGATATGACAGAAATCATGTT----- 477
Db 393 SerThrAlaCysLeuGlyGlyAlaIleAlaAlaGlnGluIleValSerIleGlnAsn 412
QY 478 ---CCTCAAGGAATT---TACGATGGGAGACGTTTAACTGCTATCATTTCCCTATGCTT 531
Db 413 AsnGlnAlaGlyIleSerPheGluGlyGlyAlaSerPheGlyGly----- 428
QY 532 ATAGGAGATCCGAGTGGACTGTTTTCGCGAGAGTAAATCAATATAAATCTT 591
Db 429 ---GlyIleAlaCysGlySerPheSerAlaGlyAlaSerValLeuGlyThrIle 447
QY 592 GAC-----AATTCATATGACGCTTTCCTTTTAAGT-----TGTTTGGGAACCTTA 636
Db 448 AspIleSerIleAsnLeuGlyAlaIleSerPheSerArgThrLeuCysThrThrSerAsp 467
QY 637 TTAGGGAGTTTACTGTTTATAGGAGGAGCACTCGTTGACTTCGAGACATACGACT 696
Db 468 LeuGlyGlnMetGluTyrGlnGly---GlyGlyAlaLeuPheGlyGluAsnIleSerLeu 486
QY 697 TCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATGAGGGT 756
Db 487 SerGluAsnAlaGlyValLeuThr----- 494
QY 757 TTTAAAGAA-----TTATCTCTTTTCCAATTCGAATTCATTACTTTCGCTACTGCCT 807
Db 495 PheIleAspAsnIleValIleThrPheAlaSerAsnGlyIleLeuGlyGlyAla 514
```

```
QY 808 GCTGCAACGACACTAATAAGGGTAGCCAGACTCCGACGACAAACATCTACACCGTCTAATGCT 867
Db 515 IleLeuAlaThrGlyLysValGluIleThrAsnAsnSerGluGlyIleSerPheThrGly 534
QY 868 ACTATTTATTTAAACAGACTCTTTGTTTACTCAATAATAAGAACTTCTCATCTTATAGT 927
Db 535 Asn-----AlaArgAlaProGlnAlaLeuProThrGlnGluGluPheProLeuPheSer 552
QY 928 AAT-----TTAGTCTCTCGA-----GATGGGGAGCTATAGATGCTAAG 966
Db 553 LysLysGluGlyArgProLeuSerSerGlyTyrSerGlyGlyGlyAlaIleLeuGlyArg 572
QY 967 AGCTTAAACGGTTCAAGGAATTAGCAAGCTTTGCTCTTCCAAAGAAATACTGCTCAAGCT 1026
Db 573 GluValAlaIleLeuHis---AsnAlaAlaValPheGluGlnAsnArgLeuGlnCys 591
QY 1027 GAT-----GGGGGAGCTTGTCAAGTAGTAGCACC 1053
Db 592 SerGluGluGluAlaThrLeuLeuGlyCysGlyGlyGlyAlaValHisGlyMetAsp 611
QY 1054 AGTTTCTCTGCTATGCTAAACGAGCTCTCTATTGCTTTGTAGCGAATCTTGCA----- 1107
Db 612 SerThrSerIleValIleValAsnSerSer---ValArgPheGlyAsnAsnTyrAlaMetGly 630
QY 1108 ---GGAGTAAGAGGGGGGATTTGCTGCTGCTTTCAGGATGGGCGAGGAGTGTCTATCA 1164
Db 631 GlnGlyValSerGlyGlyAlaLeuLeuSer----- 640
QY 1165 TCTACTTCAACAGAGATCCAGTAGTAGTATTTTTCAGAAANTACTCGCGTAGAGTTTGAT 1224
Db 641 -----LysThrValGlnLeuAlaGlyAsnGlySerValAspPheSer 654
QY 1225 GGGAACTAGCCGAGTAGGAGGAGGATTAC-----TCCTACGGGAACGTTGCTTTC 1278
Db 655 ArgAsnIleAlaSerLeuGlyGlyAlaLeuGlnAlaSerGluGlyAsnCysGluLeu 674
QY 1279 CTCAATAATAAGAAAAACCTTG----- 1299
Db 675 ValAspAsnGlyTyrValLeuPheArgAspAsnArgGlyArgValTyrGlyAlaIle 694
QY 1300 -----TTTCTCAAC 1308
Db 695 SerCysLeuArgGlyAspValValIleSerGlyAsnLysGlyArgValGluPheLysAsp 714
QY 1309 ATGTTGCTTCTCTCTTTTACATTGCTGCTAAG-----CAACCA 1347
Db 715 AsnIleAlaThrArgLeuTyrValGluGluThrValGluLysValGluValGluPro 734
QY 1348 ACAAGTGGACAGGCTTCTAATACGAGTAATAATTAC----- 1383
Db 735 AlaProGluGlnLysAspAsnAsnGluLeuSerPheLeuGlyArgAlaGluGlnSerPhe 754
QY 1384 -----GGAGAT----- 1389
Db 755 IleThrAlaAlaAsnGlnAlaLeuPheAlaSerGluAspGlyAspLeuSerProGluSer 774
QY 1390 -----GGAGGAGCTATCTTCTGT 1407
Db 775 SerIleSerSerGluLeuAlaLysArgArgGluCysAlaGlyGlyAlaIlePheAla 794
QY 1408 AAG-----AATGGTGCCCAAGCAGGATCCCAATAAC----- 1437
Db 795 LysArgValArgIleValAspAsnGlnGluAlaValPheSerAsnAsnPheSerAsp 814
QY 1438 -----TCTGATCAGTT-----TCCTTTGATGGAGAG 1464
Db 815 IleTyrGlyGlyAlaIlePheThrGlySerLeuArgGluGluAspLysLeuAspGlyGln 834
QY 1465 -----GGAGTAGTTTCTTTTACTAGCAATGATGCT 1494
Db 835 IleProGluValLeuIleSerGlyAsnAlaGlyAspValValPheSerGlyAsnSerSer 854
```


Qy 1495 GCTGGGAAA-----GGGGAGCTATTATTATGCGCAAAAAGCTCTCG 1533
Db 855 LysArgAspGluHisLeuProHisThrGlyGlyAlaIleCysThrGlnAsnLeuThr 874
Qy 1534 GTTGCT---AACTGTGGCCCTGTACAATTTTAAGGAATATCGCTAAATGATGGTGGAGCG 1590
Db 875 IleSerGlnAsnThrGlyAsnValLeuPheTyrAsnAsnValAlaCysSerGlyGlyAla 894
Qy 1591 ATTTATTATGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTTATGGAGATATTATTTTC 1650
Db 895 ValArgIleGluAspHisGlyAsnValLeuLeuGluAlaPheGlyGlyAspIleValPhe 914
Qy 1651 GATGGAACTTAAAGAACGCCAACAGAAATGCTGCCGATGTTAATGGCGTAACCTGTG 1710
Db 915 LysGlyAsnSerSerPheArgAlaGln----- 923
Qy 1711 TCCTCACAGCATTTCGATG---GGATCGGGAGGGAAATAACGACATTAAAGAGCTAAA 1767
Db 924 GlySerAspAlaIleTyrPheAlaGlyLysGluSerHisIleThrAlaLeuAsnAlaThr 943
Qy 1768 GCAGGGCATCAGATTCTCTTAATGATCCCATCGAGATGGCAAAACGGAATAAACAGCCA 1827
Db 944 GluGlyHisAlaIleValPheHisAspAlaLeuValPheGlu-----AsnLeuGluGlu 961
Qy 1828 CGCAGTCTTCCAAACTCTTAAATTAACGATGTGAA-----GGATACACAGGGGAT 1881
Db 962 ArgLysSerAlaGluValLeuLeuIleAsnSerArgGluAsnProGlyTyrThrGlySer 981
Qy 1882 ATGTGTTTGTCTAAAGGAGCAGTACTTGTACCAAAATGTTACGATAGACAGCAAGG 1941
Db 982 IleArgPheLeuGluAlaGluSerLysValProGlnCysIleHisValGlnGlnGlySer 1001
Qy 1942 ATGTGTTCTTCGAAAGCGAAAATTATCAGTGAATCTCTAAAGTCAG---ACAGGTGGG 1998
Db 1002 LeuGluLeuLeuAsnGlyAlaThrLeuCysSerTyrGlyPheLysGlnAspAlaGlyAla 1021
Qy 1999 AGTCGTATATGAAGCTGGAGTACATGGGATTTTGTGA---ACTCCACACACCAACAA 2055
Db 1022 LysLeuValLeuAlaGlyAlaLysLysIleLeuAspSerGlyThrProValGln 1041
Qy 2056 CAGCCTCTCGGCTAAAT-----CAGTTGATCAGCCTTTCCAACTGCATTG 2103
Db 1042 GlnGlyHisAlaIleSerLysProGluAlaGluLeuGluSerSerSerGluProGluGly 1061
Qy 2104 TCTCTTCTCTTCTGTAGCAACCAATGAGTACGATTCCTCCACCAATCCTCCAGCG 2163
Db 1062 AlaHisSerLeuTrpIleAlaLysAla-----GlnThrValProMet 1077
Qy 2164 CAAGATTCTCAT----- 2175
Db 1078 ValAspIleHisThrIleSerValAspLeuAlaSerPheSerSerSerGlnGlnGly 1097
Qy 2176 -----CCTGCAGTCAATGGTAGCACACTGCTGGTCTGTTTACAATATGATGGG 2223
Db 1098 ThrValGluAlaProGlnValIle-----ValProGlyGlySerTyrValArgSerGly 1115
Qy 2224 CCTATCTTTTGTAG-----GATTGGATGATACAGCTTATGATAGGTATGATGGCTA 2277
Db 1116 GluLeuAsnLeuGluValAsnThrThrGlyThrGlyTyrGluAsnHisAlaLeuLeu 1135
Qy 2278 GGTTCCTAAATAAATCAATCTCTGAAATTTACAGTTAGGAGCTAAAGCCCCAGCTAAT 2337
Db 1136 LysAsnGluAlaLysValProLeuMetSerPheValAlaSerGlyAspGluAlaSerAla 1155
Qy 2338 GCCCATCAGATTGACTCTAGGAAT----- 2367
Db 1156 GluIleSerAsnLeuSerValSerAspLeuGlnIleHisValThrProGluIleGlu 1175
Qy 2368 ATGCCTAAGTATGGCTATCAAGAGAGCTTGGCTGGGATCTCTAATACAGCAAT 2427
Db 1176 GluAspThrTyrGlyHisMetGlyAspTrpSerGluAla-----LysIleGln 1191
Qy 2428 AATGTCCTTACTCTGAAAGCTACATGAGCTAAAGCTGGGTATTAATCTCGGGCTGAG 2487

Db 1192 AspGly-----ThrLeuValIleSerTrpAsnProThrGlyTyrArgLeuAspProGln 1209
Qy 2488 CGAGTAGCTTCTTGGTTCCAAATAGTTATGG-----GGATCCATTTTA----- 2532
Db 1210 LysAlaGlyAlaLeuValPheAsnAlaLeuTrpGluGluGlyAlaValLeuSerAlaLeu 1229
Qy 2533 ---GATATACGATCTGGCATTCAGCAATTCACCAAGTGTGGATGGCGCTCTTATTGT 2589
Db 1230 LysAsnAlaArgPheAlaHisAsnLeuThrAlaGlnArgMetGlu---PheAspTyrSer 1248
Qy 2590 CGAGGATTATGGGTTTCTGGAGTTTCGAATTTCTCTCATCATGACCGCGATCTTAGGT 2649
Db 1249 ThrAsnValTrpGlyPheAlaPheGlyGlyPheArgThrLeuSerAlaGluAsnLeuVal 1268
Qy 2650 -----CAGGATATCGGTATATTAGTGGGGTATTCTCTTAGGAGCAAAC----- 2694
Db 1269 AlaIleAspGlyTyrLysGlyAlaTyrGlyAlaSerAlaGlyValAspIleGlnLeu 1288
Qy 2695 ---TCCTACTTTGGATCATCGATGTTTGGTCTAGCATTTACGAAGTATTTGGTAGACT 2751
Db 1289 MetGluAspPheValLeuGlyValSerGlyAlaAlaPheLeuGlyLysMetAspSerGln 1308
Qy 2752 AAA---GATTATGATGATCGATGTTTGGTCTAGCATTTACGAAGTATTTGGTAGACT 2808
Db 1309 LysPheAspAlaGluValSerArgLys-----GlyValValGlySerValTyrThr 1325
Qy 2809 TCTACCCCAACAGCTTTATGTGATCTCTATTGTTGGAGATGCGTTTATCCGTGCTAGC 2868
Db 1326 -----GlyPheLeuAlaGlySerTrpPhePheLysGlyGln 1337
Qy 2869 TAGGGTTTGG-----AATCAGCATATGAAAACCTCATATATACATTTTGCAGAGAGAGC 2922
Db 1338 TyrSerLeuGlyGluThrGlnAsnAspMetLysThrArgTyrGlyValLeuGlyLysSer 1357
Qy 2923 GATGTTCTGGGATAATAACTGCTGCTGCGAGAGATTGGAGCGGATACCGATTGTG 2982
Db 1358 SerAlaSerTrpThrSerArgGlyValLeuAlaAlaLeuValGluTyrArgSerLeu 1377
Qy 2983 ATTACTCCATCTAAG-----CTCTATTGTAAGTAGTTGCGCTCTCTTCGTCGAAGCTGAG 3036
Db 1378 ValGlyProValArgProThrPheTyrAlaLeuHisPheAsnProTyrValGluValSer 1397
Qy 3037 TTTTCTTATGCGCATCATGAATCTTTTACAGAGAAGCGATCAAGCTCGGCGATTCAAG 3096
Db 1398 TyrAlaSerMetLysPheProGlyPheThrGluGlnGlyArgGluAlaArgSerPheGlu 1417
Qy 3097 AGCGGACATCTCTAAATCTATCAGTCTCTGTTGGAGTGAAGTTTGATCGATGT----- 3150
Db 1418 AspAlaSerLeuThrAsnIleThrIleProLeuGlyMetLysPheGluLeuAlaPheIle 1437
Qy 3151 TCTAGTACACATCTCTAAATAATATAGCTTTTATGGCGCTTATATCTGTGATGCTTATCGC 3210
Db 1438 LysGlyGlnPheSerGluValAsnSerLeuGlyLysSerTyrAlaTrpGluAlaTyrArg 1457
Qy 3211 ACCATCTCTGTGATCGAGACCAACGCTCTCTATCCCATCAAGACATGAGCAACAGATGCC 3270
Db 1458 LysValGluGlyGlyAlaValGlnLeuLeuGluAlaGlyPheAspTrpGluGlyAlaPro 1477
Qy 3271 TTTCTATTAGCAGACATGGAGTTGTGCTT-----AGAGATCTATG 3312
Db 1478 MetAspLeuProArgGlnGluLeuArgValAlaLeuGluAsnAsnThrGluTrpSerSer 1497
Qy 3313 TATGCTTCTCTAAACAAGTAATATAGAAGTATATGGCCATGGAAGATATGAGTATCGAGAT 3372
Db 1498 TyrPheSerThrValLeuGlyLeuThrAlaPhe---CysGlyGlyPheThrSerThrAsp 1516
Qy 3373 GCTTCTCGAGGCTATGGTTGATGTCGAGGAAGTACAGTCCGGTTC 3417
Db 1517 SerLysLeuGlyTyrGluAlaAsnThrGlyLeuArgLeuIlePhe 1531
RESULT 10
US-11-018-868-18

```

; Sequence 18, Application US/11018868
; Publication No. US2006003487A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; INVENTOR: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP231152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-18

Alignment Scores:
Pred. No.: 2,52e-43 Length: 1531
Score: 597.50 Matches: 293
Percent Similarity: 36.7% Conservative: 153
Best Local Similarity: 24.1% Mismatches: 474
Query Match: 7.6% Indels: 295
DB: 11 Gaps: 54

US-10-701-844-1 (1-4435) x US-11-018-868-18 (1-1531)
QY 430 TCTTGTCTCTTAAATCGGGGGATATGCGAGAAATCATGTT----- 477
DB 393 SerThrAlaCysLeuGlyGlyAlaIleAlaGlnGluIleValSerIleGlnAsn 412
QY 478 ---CCTCAAGGAATT---TACGATGGGAGACGTTAACTGATCATCTTCCCTATACCTGTT 531
DB 413 AsnGlnAlaGlyIleSerPheGluGlyGlyLeuAlaSerPheGlyGly----- 428
QY 532 ATAGGAGATCCGAGTCGGACTACTGTTTCTCGAGGAGAGGTAACTTAAATAATCTT 591
DB 429 ---GlyIleAlaCysGlySerPheSerAlaGlyGlyAlaSerValLeuGlyThrIle 447
QY 592 GAC-----AATCTATTGCGCTTTCCTTTAAGT-----TGTTTGGGAACCTTA 636
DB 448 AspIleSerIleAsnLeuGlyAlaIleSerPheSerArgThrLeuCysThrThrSerAsp 467
QY 637 TTAGGAGATTTTACTGTTTATAGGAGGACACTCGTTGACTTTTCGAGACATACGAGCT 696
DB 468 LeuGlyGlnMetGluThrGlnGly---GlyGlyAlaLeuPheGlyGluAsnIleSerLeu 486
QY 697 TCTACAAATGGGCGACCTTAAGTAATAGCGCTGCTGATGCGACTGTTTACTATTGAGGGT 756
DB 487 SerGluAsnAlaGlyValLeuThr-----TGTTTGGGAACCTTA 494
QY 757 TTTAAAGAA-----TTATCTTTTCCAAATGCAATTCATTACTTGCCTGCTGCT 807
DB 495 PheIleAspAsnIleValIleThrPheAlaSerAsnGlyIleLeuGlyGlyAla 514
QY 808 GCTGCAACGACTAATAGGGTAGCCAGACTCCGACGACACATCTACACCGTCTAATGGT 867
DB 515 IleLeuAlaThrGlyIleValGluIleThrAsnAsnSerGluGlyIleSerPheThrGly 534
QY 868 ACTATTATTCTAAACAGACTCTTTTGTATTACTCAATTAATGAGAGTTCTCATTTCTATAGT 927
DB 535 Asn-----AlaArgAlaProGlnAlaLeuProThrGlnGluPheProLeuPheSer 552
QY 928 AAT-----TTAGTCTCTGGA-----CATGGGAGCTATAGATGCTAAG 966
DB 553 LysLysGluGlyArgProLeuSerSerGlyTy-SerGlyGlyAlaIleLeuGlyArg 572
QY 967 AGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCTCAAGAAAATACTGCTCAAGCT 1026
DB 573 GluValAlaIleLeuHis---AsnAlaAlaValPheGluGlnAsnArgLeuGlnCys 591

```

```

QY 1027 GAT-----GGGGAGCTTGTCAAGTAGTCACC 1053
DB 592 SerGluGluGluAlaThrLeuLeuGlyCysCysGlyGlyAlaValHisGlyMetAsp 611
QY 1054 AGTTTCTCTGTATGGCTTAAGAGCTCCTATTGCTTTGTAGCGAATGTGCA----- 1107
DB 612 SerThrSerIleValGlyAsnSerSer---ValArgPheGlyAsnAsnTyrAlaMetGly 630
QY 1108 ---GGAGTAAGGGGGAGGATTGCTGCTGTTTCAGGATGGCGACAGGAGTGTCA 1164
DB 631 GlnGlyValSerGlyGlyAlaLeuLeuSer----- 640
QY 1165 TCTACTTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATPACTGCGTAGAGTTTGTAT 1224
DB 641 -----LysThrValGlnLeuAlaGlyAsnGlySerValAspPheSer 654
QY 1225 GGGAGCTAGCCCGAGTAGGAGGAGGATTTAC-----TCTCAGGAGACCTGTTGTTTC 1278
DB 655 ArgAsnIleAlaSerLeuGlyGlyAlaLeuGlnAlaSerGluGlyAsnCysGluLeu 674
QY 1279 CTGAATAATGGAAAAACCTTG----- 1299
DB 675 ValAspAsnGlyTyrValLeuPheArgAspAsnArgValTyArgValTyArgValTyArgVal 694
QY 1300 -----TTTCTCAAC 1308
DB 695 SerCysLeuArgGlyAspValIleSerGlyAsnLysGlyArgValGluGluPheLysAsp 714
QY 1309 AATGTTGCTTCTCTGTTTACATTGCTGCTAAG-----CAACCA 1347
DB 715 AsnIleAlaThrArgLeuTyrValGluGluThrValGluLysValGluGluValGluPro 734
QY 1348 ACAAGTGGACAGCTTCTTAATACGAGTAATATTAC----- 1383
DB 735 AlaProGluGlnLysAspAsnAsnGluLeuSerPheLeuGlyArgAlaGluGlnSerPhe 754
QY 1384 -----GGAGAT----- 1389
DB 755 IleThrAlaAlaAsnGlnAlaLeuPheAlaSerGluAspGlyAspLeuSerProGluSer 774
QY 1390 -----GGAGGACTATCTTCTGT 1407
DB 775 SerIleSerSerGluGluLeuAlaLysArgGluCysAlaGlyAlaIlePheAla 794
QY 1408 AAG-----AATGTCGCAAGCAGGATCCCAATAC----- 1437
DB 795 LysArgValArgIleValAspAsnGlnGluAlaValPheSerAsnAsnPheSerAsp 814
QY 1438 -----TCTGATCAGTT-----TCCTTTGATGAGGAG 1464
DB 815 IleTyArgGlyAlaIlePheThrGlySerLeuArgGluGluAspLysLeuAspGlyGln 834
QY 1465 -----GGAGTAGTTTCTTTTAGTATGATGAGTATGAGT 1494
DB 835 IleProGluValLeuIleSerGlyAsnAlaGlyAspValPheSerGlyAsnSerSer 854
QY 1495 GCTGGGAAA-----GGGGAGCTATTATGCCAAAAGCTCTCG 1533
DB 855 LysArgAspGluHisLeuProHisThrGlyGlyAlaIleCysThrGlnAsnLeuThr 874
QY 1534 GTTGCT---AAGTGTGCGCTGTACAAATTTTAAAGGAATATCGCTAATGATGTTGAGCG 1590
DB 875 IleSerGlnAsnThrGlyAsnValLeuPheTyArgAsnValAlaCysSerGlyGlyAla 894
QY 1591 ATTTATTAGGAAATCTGGAGAGCTCAGTTTATCTGCTGATTTATGAGATATATTTTC 1650
DB 895 ValArgIleGluAspHisGlyAsnValLeuLeuGluAlaPheGlyGlyAspIleValPhe 914
QY 1651 GATGGGAATCTTAAAGAACACAGCAAGAGATGTCGCCGATGTTAATGCGCTAACTGTG 1710
DB 915 LysGlyAsnSerPheArgAlaGln----- 923

```



```

; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 975
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-47

Alignment Scores:
Pred. No.: 2,356-39 Length: 975
Score: 551.00 Matches: 279
Percent Similarity: 38.1% Conservative: 148
Best Local Similarity: 24.9% Mismatches: 405
Query Match: 7.0% Indels: 290
DB: 11 Gaps: 57

US-10-701-844-1 (1-4435) x US-11-103-957-47 (1-975)
Qy 388 ACCTCTTCCATAGTCTTCTTCAATGATTCTAGCTATTCTTGCTGCTTTAAAT 447
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 32 ThrLeuIleProLysPheLeuLeuGlyAlaLeuIleValTyrAlaProTyrSerPhe--- 50
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 448 GGGGGGGATATGCAGCAGAAATCATGGTT-----CCTCAAGGAATTTAC 492
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 51 -----AlaGluMetGluLeuAlaIleSerGlyHisGlnGly---Lys 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 493 GATGGGAGACGTTAACTGTATCATTTCCCTATCTGTATTAGGAGATCCGAGTGGACT 552
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 65 AspArgAspThrPheThrMet-----IleSerSerCysProGluGlyThr 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 553 ACTGTTTTTCTGACGAGGAGTTACATTAATA-----AATCTGCACATTTCTATTGCA 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 80 AenTyrIleIleAsnArgLysLeuLeuSerAspPheSerLeuLeuAsnLysValSer 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 607 GCTTTGCTTTAAAGTTGTTGGGAACATTATTAGGAGTTTACTTTTAGGGAGAGA 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 100 Ser-----GlyGlyAlaPheArgAsnLeuAlaGlyLysIleSerPheLeuGlyLysAsn 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 667 CAC-----TCGTTGACTTTCGAGAACATACCGACTTCTACAAATGGGGCAGCTCTAAGT 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 118 SerSerAlaSerIleHisPheLysHisIleAsnIleAsnGlyPheGlyAlaGlyValPhe 137
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 ATACGGCTGCTGAGGACTGTTTACTATTAGGGGTTTAAAGAATTATCTTTTCCAT 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 138 SerGluSerSer-----IleGluPheThrAsp 146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 781 TGCATTTCAATTAATCTGCGTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCG 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 LeuArgLysLeuValAlaPhe-----GlySerGluSer--- 157
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 841 ACACAAATCTACACCGTCTAATGGTACTATTATTTCTTAAACAGACTCTTTTGTACTC 900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 158 -----ThrGlyLysIlePheThrAlaLysGluAspIleSerPheLys 171
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 901 AATAATGAGAAGTCTCATCTTATAGTAATAGTCTCTCGAGATGGGGAGCTATAGAT 960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 172 AsnAsnHisIleAlaPheArgAsnAsnIleThrLysGlyAsnGlyValIleGln 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 961 GCTAAGAGCTTAACGGTTCAAGGAATTAGCAAG----- 993
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 192 -----LeuGlnGlyAspMetLysGlySerValSerPheValAspGlnArg 206
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 994 ---CTTTGTGCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTTGTCAAGTAGTC 1050
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 207 GlyAlaIleIlePheThrAsnAn-----GlnAlaVal 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1051 ACCAGTTCTCTGCTATGCTAACGAGGCTCTTATGCTTTGTAGCGAATGTTGCAGGA 1110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 218 ThrSer---SerSerMetLysHisSerGly----- 226
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1111 GTAAGGGGGGGAGGATTCTGCTGTTTCAGGATGGGCAGCAGGAGTGTCTATCTACT 1170
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```
QY 2563 GCAAGTGTGGAGCGCTCTATTGTGCGAGATTATGGTTCTCGAGATT----- 2613
Db      |||||
699 ArgSerLeuIleProThrSerTyrPhe---GlyValLeuIleGlyGlyThrGlyAlaGlu 717
2614 -----TCGAAATTCCTCTCATCATGACCGCGAGTCTTTAGGTCAG 2652
718 MetSerThrHisSerSerGluGluSerPheIleSerArgLeuGlyAlaThrGlyThr 737
2653 GGAATATCGTATATTAGTGGGGTTTATTCCTTAGGAGCAAACTCCTATTGGTCATCG 2712
Db      |||||
738 SerIleIleArgLeuThrProSerLeuThrLeuSerGlyGly-----GlySerHis 754
2713 ATGTTTGGCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTATGTAGTGTGCT 2772
Db      |||||
755 MetPheGlyAspSerPhe-----ValAlaAspLeu 764
2773 TCCAAATCATCATCTGTCATAGGATCCGTTTAT----- 2805
Db      |||||
765 ProGluHisIleThrSerGluGlyIleValGlnAenValGlyLeuThrHisValTrpGly 784
2806 ---CTATCTACCAACAAGCTTTATGTGATCTCTATTGTTGGAGATGCGTTTATCGT 2862
Db      |||||
785 ProIleThrValAenSerThrLeuCysAlaIle---LeuAspHisAenAlaMetValArg 803
2863 GCTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAGAGC 2922
Db      |||||
804 IleCys-----SerIlyAspHisThrTyrGly----- 813
2923 GATGTTCTGGGATAAATACGTCTGCTGGAGAGATTGGAGCGGATACCGATTGTG 2982
Db      |||||
814 ---LysTrpAspThrPheGlyMetArgGlyThrLeuGlyAlaSerTyrThrPheLeu 831
2983 ATTACTCCATCTAAGCTCTATTGTAATGAGTTGGCTCTCTTCGTCGAAGCTGAGTTTCT 3042
Db      |||||
832 -----GluTyrAspGlnThrMetArg-----ValPheSer 841
3043 TATGCCGATCATGAA-----TCTTTTACAGAGAGGCGATCAA 3081
Db      |||||
842 PheAlaAenIleGluAlaThrAenIleLeuGlnAArgAlaPheThrGluThrGlyTyrAen 861
3082 GCTGGGATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGAGTGAAGTTT 3141
Db      |||||
862 ProArgSerPheSerLysThrLysLeuAenIleAlaIleProIleGlyGlyTyr 881
3142 GATCGATGTTCTAGTACATCTCTATAAATATAGCTTTATGCGC-----GCT 3189
Db      |||||
882 GluPheCysLeuGlyAen-----SerSerPheAlaLeuLeuGlyLysGlySerIleGly 899
3190 TATATCTGTGATGCTTATGCGACCATCTCTGTGACTGAGACACGCTCTATCCCATCAA 3249
Db      |||||
900 TyrSerArgAspIleLysArgGluAenProSerThrLeuAlaHisLeuAlaMetAenAsp 919
3250 GAGACATGAGCAACAGAT-----GCCCTTTCATTTAGCAAGACATGA 3291
Db      |||||
920 PheAlaTrpThrThrAenGlyCysSerValProThrSerAlaHisThrLeuAlaAenGln 939
3292 GTTGGTGTAGGATCTATGATCTCTCTTAACAAGTAATATAGAAGTATATGCCAT 3351
Db      |||||
940 LeuIleLeuArg-----TyrLysAlaCysSerLeuTyrIleThrAlaTyrThrIle 956
3352 GGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTGGTGGCAGGAGTAGTAGTC 3411
Db      |||||
957 AenArg---GluGlyLysAenLeuSerAen-----SerLeuSerCysGlyGlyTyrVal 973
3412 CGGTTTC 3417
Db      |||||
974 GlyPhe 975
```

RESULT 13

US-11-018-868-50

; Sequence 50, Application US/11018868

; Publication NO. US20060034871A1

```
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (P23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50
; LENGTH: 975
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-50
```

```
Alignment Scores:
Pred. No.: 2,35e-39 Length: 975
Score: 551.00 Matches: 279
Percent Similarity: 38.1% Conservative: 148
Best Local Similarity: 24.9% Mismatches: 405
Query Match: 7.0% Indels: 290
DB: 11 Gaps: 57
```

US-10-701-844-1 (1-4435) x US-11-018-868-50 (1-975)

```
QY 388 ACGTCTTCCATAAGTTCTTTTCAATGATTCCTAGCTTATTCTGCTGCTCTTTAAAT 447
Db      |||||
32 ThrLeuIleProLysPheLeuLeuGlyAlaLeuIleValTyrAlaProTyrSerPhe--- 50
448 GGGGGGGGATATGACGACAGAAATCATGTT-----CCTCAAGGAATTTAC 492
Db      |||||
51 -----AlaGluMetGluLeuAlaIleSerGlyHisLysGlnGly---Lys 64
493 GATGGGAGAGCGTTAATCTGTATCATTTCCCTATCTATCTGTATAGGAGATCCGAGTGGACT 552
Db      |||||
65 AspArgAspThrPheThrMet-----IleSerSerCysProGluGlyThr 79
553 ACTGTTTTTCTGCGAGGAGAGTTAACTTAACTTAAAA-----AATCTTGACAAATCTTATTCGA 606
Db      |||||
80 AsnTyrIleIleAenArgLysLeuLeuSerAspPheSerLeuLeuAenLysValSer 99
607 GCTTTGCCCTTTAAGTTGTTTGGGAACCTATTAGGAGGTTTACTGTTTATAGGGAGAGGA 666
Db      |||||
100 Ser-----GlyGlyAlaPheArgAenLeuAlaGlyLysIleSerPheLeuGlyLysAen 117
667 CAC-----TCGTTGACCTTTCGAGAACATACGGACTTCTACAAATGGGGGAGCTTAAGT 720
Db      |||||
118 SerSerAlaSerIleHisPheLysHisIleAenIleAenGlyPheGlyAlaGlyValPhe 137
721 AATAGCGCTGCTGATGGACTGTTTACTATTGAGGGTTTTTAAAGAATTATCTTTTCCCAAT 780
Db      |||||
138 SerGluSerSer-----IleGluPheThrAsp 146
781 TGCAATTCATTACTTTCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCG 840
Db      |||||
147 LeuArgLysLeuValAlaPhe-----GlySerGluSer--- 157
841 ACGACAACTCTACACCGCTCTAATGGTACTATTATTCTTAAACAGATCTTTTGTACTCTC 900
Db      |||||
158 -----ThrGlyGlyIlePheThrAlaLysGluAspIleSerPheLys 171
901 AATATGAGAAGTTCTCAATCTATAGTAATTAGTCTCTGAGAGATGGGGGAGCTATAGAT 960
Db      |||||
172 AenAenHisHisIleAlaPheArgAenAenIleThrLysGlyAenGlyGlyValIleGln 191
961 GCTAAGAGCTTAAACGGTTTCAAGGAATTTAGCAAG----- 993
Db      |||||
192 -----LeuGlnGlyAspMetLysGlySerValSerPheValAspGlnArg 206
994 ---CTTTGTGTTCTTCCAAGAAAATACTGCTCAAGCTGATGGGGGAGCTTGTCAAGTAGTC 1050
```

```

Db      207 GlyAlaIlePheThrAsnAn-----GlnAlaVal 217
QY      1051 ACCAGTTCTCTGCTATGCTAACAGAGGCTCTATTATGCTTTGTAGCGAATGTCAGGA 1110
Db      218 ThrSer---SerSerMetIysHisSerGly----- 226
QY      1111 GTAAGAGGGGAGGATTCCTGCTGTTCCAGGATGGCAGCAGGAGTGTCATCTACT 1170
Db      227 ---ArgGlyGlyAlaIleSerGlyAspPheAlaGlySerArg----- 239
QY      1171 TCAACAGAGATCCAGTACGATAGTTTTCAGAAATACCTGCGGTAGAGTTTGATGGGAAC 1230
Db      240 -----IleLeuPheLeuAsnAsnGlnGlnIleThrPheGluGlyAsn 253
QY      1231 GTAGCCCGAGTAGGAGAGGATTTACTCTCTCTACCGGAACGTGCTTCTCTGTAATAATGA 1290
Db      254 SerAlaValHisGlyGlyAlaIleIleThrAsn-----LysAsnGly 266
QY      1291 AAMACCTTGTCTCAACATGTTGCTTCTCTGTTTACATGCTGCTAGCAACCAACA 1350
Db      267 LeuValGluPheLeuGlyAsn---AlaGlyPro-----LeuAlaPheLysGluAsnThr 283
QY      1351 -----AGTCGACAGGCTCTAATACAGTAAT-----AATTACGGAGATGGA 1392
Db      284 ThrIleAlaAsnGlyGlyAlaIleIleThrSerAsnPhelysAlaAsnGlnIleThrSer 303
QY      1393 GGAAGCTATCTCTGTAAGATGGTGCACAGCAGGATCCAACTCTGATCAGTTTCC 1452
Db      304 ProIleLeuPheSerGlnAsnHisAla-----AsnLysLysGlyAlaIle 319
QY      1453 TTT-----GATGGAGAGGAGTAGTTTCTTTAGTACCAATGTA 1491
Db      320 TyrAlaGlnIleValAsnLeuGluAsnGlnAspThrIleArgPheGluLysAsnThr 339
QY      1492 GCTGCTGGGAAAGGGAGCTATTATTGCCCCAAAGCTCTCGTGTGCTAACTGTGGCCCT 1551
Db      340 AlalysGluGlyGlyAlaIleIleThrSerSerGlnCysSerIleThrAlaHisAsnThr 359
QY      1552 GTACAAATTTTAAGGAATATCGCTAATGAT-----GGTGGAGGATTTATTTA---GGA 1602
Db      360 IleIlePheSerAspAsnAlaAlaGlyAspLeuGlyGlyGlyAlaIleLeuLeuGluGly 379
QY      1603 GAATCTGGAGAGCTCAGTTTATCTGCTGATATGAGATATTTATTTTCGATGGGAAT--- 1659
Db      380 LysLysProSerLeuThrLeuIleAlaHisSerGlyAsnIleAlaPheSerGlyAsnThr 399
QY      1660 ---CTTAAAGAACAGCCAAAGAGATGCTCCGATGTTAATGGCGTAACGTCCTCA 1716
Db      400 MetLeuHisIleThrLysLysAlaSerLeuAspArgHisAsnSerIleLeuIleLysGlu 419
QY      1717 CAAGCATTTTCGATGGATCGGAGGAGGAAATAACGACATTAAGCTAAAGCAGGCGAT 1776
Db      420 AlaProTyrLysIle-----GlnLeuAlaAlaAsnLysAsnHis 432
QY      1777 CAGATTTCTTTAATGATCCCATCGAGATGCGCAACCGAAATACCGAGCCGCGCTCT 1836
Db      433 SerIleHisPheAspProValMetAlaLeuSerAlaSerSerProIleGln--- 451
QY      1837 TCCAACTCTTAAATTAACGATGCTCAGGATACACA-----GGG 1878
Db      452 -----IleAsnAlaProGluTyrGluThrProPhePheSerProLysGly 466
QY      1879 GATATTGTTTT-----GCTAATGGA 1899
Db      467 MetIleValPheSerGlyAlaAsnLeuLeuAspAlaArgGluAspValAlaAsnArg 486
QY      1900 AGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGAGGATGTTCTTCGTGAAAG 1959
Db      487 ThrSerIlePheAsnGlnProValHisLeuTyrAsnGlyThrLeuSerIleGluAsnGly 506
QY      1960 GCAAAATATCAGTAATCTCTAAGTCAGACAGTGGGATCTGTATATGGAAGCTGGG 2019

```

```

507 AlaHisLeuIleValGlnSerPheLysGlnThrGlyGlyArgIleSerLeuSerProGly 526
2020 AGTACATCGGATTTTGTAACTCCACAA-----CCACCACAACAG 2058
527 SerSerLeuAlaLeuTyrThrMetAsnSerPhePheHisGlyAsnIleSerLysGlu 546
2059 CCTCTCGCGCTAAATCAGTTGATCAGGCTTTCCAAATCTGCATTTGTCTCTTCTCTTTG 2118
547 ProLeuGluIleAsnGlyLeu-----SerPheGlyValAspIleSerProSerAsnLeu 564
2119 TTAGCAAAATGCAATGCTAGTACGATCTCTCTACCAATCTCCAGCGCAAGATTCTCATCT 2178
565 GlnAlaGluIleArgAlaGlyAsnAlaProLeuArg----- 576
2179 GCAGTCATTGGTAGCACAACTGCTGTTCTGTACAAATTAGTGGGCT-----ATCTTT 2232
577 -----LeuSerGlySerProSerIleHisAspProGluGlyLeuPhe 590
2233 TTTGAGGATTTGGATGATACAGCT-----TATGATAGTATGATTTGGCTAGTTCTAAT 2286
591 TyrGluAsnArgAspThrAlaAlaSerProTyrGlnMetGluIleLeuLeuThrSerAsp 610
2287 CAAAAAATCAATGCTCTGAAATACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCATCA 2346
611 LysIleValAspIleSerLysPhe-----ThrThr 620
2347 GATTGACTCTAGGGAATGAGATGCTAAGTATGGCTATCAAGAAAGCTGGAAGCTTGC 2406
621 AspSerLeuValThrAsnLys-----GlnSerGlyPheGlnGlyAlaIleHisPheSer 638
2407 TGGGATCTTAATACACAAATAATGCTCTTAT---ACTCTGAAGCTACATGAGCTAAA 2463
639 TrpGlnProAsnThrIleAsnAsnThrLysGlnLysIleLeuArgAlaSerTrpLeuPro 658
2464 ACTGGG---TATAATCCTGGGCTGAGCAGTAGCTCTTCTGTTCCAAATAGTTATG 2520
659 ThrGlyGluTyrValLeuGluSerAsnArgValGlyArgAlaValProAsnSerLeuTrp 678
2521 GGATCCATTATAGATATACGATCTGCG-----CATTCAGCAATTCAA 2562
679 SerThrPheLeuLeuLeuGlnThrAlaSerHisAsnLeuGlyAspHisLeuCysAsnAsn 698
2563 GCAAGTGTGAGTGGGCGCTCTTATGTCGAGGATTAATGGGTTTCTGAGTT----- 2613
699 ArgSerLeuIleProThrSerTyrPhe---GlyValLeuIleGlyThrGlyAlaGlu 717
2614 -----TCGAATTTCTTATCATCATGCCGCGATGCTTAGTCTCAG 2652
718 MetSerThrHisSerSerGluGluSerPheIleSerArgLeuGlyAlaThrGlyThr 737
2653 GGATATCGGTATATTAGTGGGCTTATCTCTTAGGAGCAAACTCTACTTTGGATCATCG 2712
738 SerIleAlaArgLeuThrProSerLeuThrLeuSerGlyGly-----GlySerHis 754
2713 ATGTTTGGTCTAGCATTTTACCAGATATTGTTAGTATTAAGATTATGATGTGTGTCGT 2772
755 MetPheGlyAspSerPhe-----ValAlaAspLeu 764
2773 TCCAATCATCATGCTTGCATAGATCCGTTTAT----- 2805
765 ProGluHisIleThrSerGluGlyIleValGlnAsnValGlyLeuThrHisValTrpGly 784
2806 ---CTATCTACCCCAACAGCTTTATGAGTCTCTATTTGTTCCGAGATGCTTTATCCGT 2862
785 ProLeuThrValAsnSerThrLeuCysAlaAla---LeuAspHisAsnAlaMetValArg 803
2863 GCTAGTACGGGTTTGGGAATCAGCATATCAAAACCTCATATATCATTTGCAAGGAGC 2922
804 IleCys-----SerLysLysAspHisThrTyrGly----- 813
2923 GATGTTCTGTTGGGATAATACTGCTGGCTGGAGAGATTGGAGCGGAGTACCGATTGTG 2982
814 -----LysTrpAspThrPheGlyMetArgGlyThrLeuGlyAlaSerTyrThrPheLeu 831

```

```
QY 2983 ATTACTCCATTAAGCTCTATTGTAATGAGTTGGCTCTTTGTCGCAAGCTGAGTTTCT 3042
Db 832 -----GlutyrAspGlnThrMetArg-----ValPheSer 841
QY 3043 TATGCCGATCATGA-----TCCTTTACAGAGGAGCGCATCAA 3081
Db 842 PheAlaAenIleGluAlaThrAenIleLeuGlnArgAlaPheThrGluThrGlyTyrAen 861
QY 3082 GCTCGGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTT 3141
Db 862 ProArgSerPheSerLysThrLysLeuAenIleAlaIleProIleGlyIleGlyTyr 881
QY 3142 GATCATGTTCTAGTACATCCTTAATAATATAGCTTTATGGCG-----GCT 3189
Db 882 GluPheCysLeuGlyAen-----SerSerPheAlaLeuLeuGlyLysGlySerIleGly 899
QY 3190 TATATCTGTGATGCTTATCGCACCATCTCTGTGTACTGAGACAACGCTCTATCCCATCAA 3249
Db 900 TyrSerArgAepIleLysArgGluAenProSerThrLeuAlaHisLeuAlaMetAenAep 919
QY 3250 GAGACATGACACACAGAT-----GCCTTTTCATTTAGCAAGACATCGA 3291
Db 920 PheAlaTyrThrAenGlyCysSerValProThrSerAlaHisThrLeuAlaAenGln 939
QY 3292 GTTGTGGTTAGAGATCTATGATGCTCTCTAACAGTAATATAGAAGTATATGGCCAT 3351
Db 940 LeuIleLeuArg-----TyrLysAlaCysSerLeuTyrIleThrAlaTyrThrIle 956
QY 3352 GGAAGATATGATATGAGATGCTTCTCGAGGCTATGTTTCAGTGCAGGAAGTAGAGTC 3411
Db 957 AenArg-----GluGlyLysAenLeuSerAen-----SerLeuSerCysGlyGlyTyrVal 973
QY 3412 CGGTTT 3417
Db 974 GlyPhe 975
```

RESULT 14

```
US-11-103-957-11
; Sequence 11, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Bertinet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; PRIOR FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-11
```

```
Alignment Scores:
Pred. No.: 7,41e-37 Length: 1034
Score: 522.50 Matches: 274
Percent Similarity: 37.1% Conservative: 157
Best Local Similarity: 23.6% Mismatches: 435
Query Match: 6.6% Indels: 295
DB: 11 Gaps: 53
```

US-10-701-844-1 (1-4435) x US-11-103-957-11 (1-1034)

```
QY 421 CTAGCTTATTTCTGCTGCTCTTTAAATGGGGGGGATATGACGACGAAATCATGTTCTCT 480
Db 7 LeuSerPheAlaCysLeuSerPhe-----PheTyrLeuSerThrIleSerIleLeu 23
QY 481 CAAGGAATTTACGATGGGAGACGCTTAACCTGTATCATTT----- 519
Db 24 GlnAla-----AenGluThrAspThrLeuGlnPheArgPheThrPheSerAep 40
QY 520 ---CCCTATATCTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTGCGAGGAGTTA 576
Db 41 ArgGluIleGlnPheValLeuAepProAla-----SerLeuIleThrAlaGlnAenIle 58
QY 577 ACATTAAATAATCTTGACAAATCTATTCCAGCTTTGCCCTTTAAGTTGTTGGGAACCTA 636
Db 59 ValLeuSerAenLeuGlnSer-----AenGly 67
QY 637 TTAGGGAGTTTTTACTGTTTTTAAAGGAGGAGCACACTCGTTAGACTTTTCGAGAACATACGACT 696
Db 68 ThrGlyAlaCysThrIleSerGlyAenThrGlnThrGlnIlePheSerAen----- 84
QY 697 TCTACAAATGGGCGAGCTCTTAAGTAATAGCCGCTGCTGATGAGCTG-----TTTACT 747
Db 85 SerValAenThrThrAlaAepSerGlyGlyAlaPheAepMetValThrThrSerPheThr 104
QY 748 ATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTTACTTCCGCTACTGCT 807
Db 105 AlaSerAepAenAlaAenLeuPhe-----CysAenAenTyr----- 117
QY 808 GCTCAACGACTAATAAGGCTAGCCAGACTCCGACGACACATCTACACCGCTTAATGCT 867
Db 118 ---CysThrHisAenLysGly-----GlyGly 125
QY 868 ACTATTTATTTCTAAACAGATCTTTTGTATCTCAATAATAGAAAGTTCTCTATCTATAGT 927
Db 126 AlaIleArgSerGlyGlyProIleArgPheLeuAenAenGlnAepValLeuPheTyrAen 145
QY 928 AATTAGTCTCTGGA-----GATGGG 948
Db 146 AenIleSerAlaGlyAlaLysTyrValGlyThrGlyAspHisAenGluLysAenArgGly 165
QY 949 GGAGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTTCCAA 1008
Db 166 GlyAlaLeuTyrAlaThrThrIleThrLeuThrGly---AenArgThrLeuAlaPheIle 184
QY 1009 GAAATACTGCTCAAGCTGATGGGAGGCTTGTCAAGTAGTCACCAAGTTTCTCTGCT--- 1065
Db 185 AenAenMetSerGlyAepCysGlyGlyAlaIleSerAlaAepThrGlnIleSerIleThr 204
QY 1066 -----ATGGCTAACGAGGCTCTATTGCC 1089
Db 205 AspThrValLysGlyIleLeuPheGluAenAenHisThrLeuAenHisIleProTyrThr 224
QY 1090 TTTGTAGCGAATGTTGCAGAGTAGTAAGAGGGAGGAGGATTGCTGCTGTTCCAGATGGCAG 1149
Db 225 GlnAlaGluAenMetAla-----ArgGlyGlyAlaIleCysSerArgAepAep----- 240
QY 1150 CAGGAGGTGTCATCTACTTCTCAACAGAAAGTCCAGTAGTAAGTTTTCAGAAATACT 1209
Db 241 -----LeuCysSerIleSerAenAenSerGlyProIleVal----- 252
QY 1210 GCGGTAGAGTTTGTAGGGAAACGTACCGAGTAGGAGGAGGAGGATTACTCTACGGGAAC 1269
Db 253 -----PheAenTyrAenGlnGlyGlyGlyGlyGlyAlaIleSerAlaThrArgCys 269
QY 1270 GTTGTCTTTCTGAATAATGGAAAAACCTTGTGTTCTCAACAATGTTGCTTCTCTCTGTTTAC 1329
Db 270 ValIleAepAenAenLysGluArgIleIlePheSerAenAen----- 283
QY 1330 ATTGCTGCTAAGCAACCAACAGTGGACAGGCTTCTTAATACGAGTAATTAATTCAGGAGAT 1389
Db 284 -----SerSerLeuGlyTyrSerGlnSerSerSerAlaSerAen 296
```

1390 CGAGGAGCTATCTTCTGTGAAGAAATCGTGGCGAAGGATCCAAATAACTCTGGATCAGTT 1449
Db |||||
297 GlyGlyAlaIleGlnThrThrGlnGlyPheThrLeu---ArgAsnAsnLysGlySerIle 315
Qy |||||
1450 TCCCTTTGAGGAGGAGGAGTACTTTCTTTAGTACCAATGTAGCTGCTGGGAAGGGGA 1509
Db |||||
316 TyrPheAsp-----SerAsnThrAlaThrHisAlaGlyGly 327
Db |||||
1510 GCTATTATGCCAAAGAGCTCTCGGTGTGTAATCTGGCCCTGTACAAATTTTAAAGTAAT 1569
Qy |||||
328 AlaIleAsnCyGlyTyriLeAspIleArgAsnGlyProValTyrPheLeuAsnAen 347
Db |||||
1570 ATCGCTAATGAGTGGCGGAGTATTATTAGAGAA-----TCTGGAGAGCTCAGTTTAA 1623
Qy |||||
348 SerAlaAlaTrpGlyAlaAlaPheAsnLeuSerLysProArgSerAlaThrAsnTyriLe 367
Db |||||
1624 TCTGCTGATATGAGATATTATTTTC-----GATGGAAAT 1659
Qy |||||
368 HisThrGlyThrGlyAspIleValPheAsnAsnAsnValValPheThrLeuAspGlyAen 387
Db |||||
1660 CTTAAAGAACACCCAAAGAGATGCTGCCGATGTGTAATGGCGTAACCTGTGCTCCTCACAA 1719
Qy |||||
388 Leu-----LeuGlyLysArgLysLeuPheHisIleAsnAsnAsnGluIleThrProTyr 405
Db |||||
1720 GCATTTTCGATCGGAGCGGAGGAAATAAAGACATTAAAGAGCTAAAGCAGGCGCATCAG 1779
Qy |||||
406 ThrLeuSerLeuGly-----AlaLysLysAspThrArg 416
Db |||||
1780 ATTCTCTTTAATGAT-----CCCATCGAGATGGCAACCGGAATAACACAG 1824
Qy |||||
417 IleTyrPheTyrAspLeuPheGlnTrpGluArgValLysGluAsnThrSerAsnAsnPro 436
Db |||||
1825 CACGCCAGCTTCCAAA---CTTCTAAATAATTAAGATGTCGTAAGATACACAGGGGAT 1881
Qy |||||
437 ProSerProThrSerArgAsnThrIleThrValAsnProGluThrGluPheSerGlyAla 456
Db |||||
1882 ATTGTTTTCCT---AATGGAACAGTACTTTGTACCAAAATGTT----- 1923
Qy |||||
457 ValValPheSerTyrAsnGlnMetSerSerAspIleArgThrLeuMetGlyLysGluHis 476
Db |||||
1924 -----ACGATAGACGAAGGAGGATTTGTTCTTCGTGAAAG 1959
Qy |||||
477 AsnTyrIleLysGluAlaProThrThrLeuLysPheGlyThrLeuAlaIleGluAspAsp 496
Db |||||
1960 GCAAAATTAATCAGTAATCTCTAAGTCAGACAGGTTGGAGT-----CTGTATATG 2010
Qy |||||
497 AlaGluLeuGluIlePheAsnIleProPheThrGlnAsnProThrSerLeuLeuAlaLeu 516
Db |||||
2011 GAAGCTGGGAGTACATGGGATTTTGTAACTCCACACCACACAGCCTCCTCGCGCT 2070
Qy |||||
517 GlySerGlyAlaThrLeuThrVal-----GlyLys 526
Db |||||
2071 AATCAGTTGATCAGCTTCCAAATCTCTGATTTGCTCTTCTTCTTTGTTAGCAAAAT 2130
Qy |||||
527 HisGlyLysLeuAsnIleThrAsnLeuGlyValIleLeuProIleIleLysGluGly 546
Db |||||
2131 GCAGTTACGAATCTCTCT-----ACCAATCTCCAGCGGCAAGATTTCTCATCTGCA 2181
Qy |||||
547 -----LysSerProProCysIleArgValAsnProGlnAspMetThrGlnAsnThrGly 564
Db |||||
2182 GTCATGCTAGCACAACTGCTGTTCTGTACAAATTTAGTGGCCCTACTCTTTT----- 2235
Qy |||||
565 Thr-----GlyGlnThrProSerSerThrSerIleSerThrProMetIleIlePheAsn 583
Db |||||
2236 -----GAGATTTTGGATGATACAGCTTATGATAGTAGTATGTTGGCTA 2277
Qy |||||
584 GlyArgLeuSerIleValAspGluAsnTyrGluSerValTyrAspSerMetAspLeu--- 602
Db |||||
2278 GGTTCATAAATAAATCAATGCTCTGAATTTACAGTTAGGAGTAAAGCCCCAGCTAAT 2337
Qy |||||
603 ---SerArgGlyLysAlaGluGlnLeuIleLeuSerIleGluThrThrAsnAspGlyGln 621
Db |||||
2338 GCCCCATCAGAT-----TTGACTCTAGGGAATCAGATG-----CCTAAGTATGGC 2382
Qy |||||

622 LeuAspSerAsnTrpGlnSerSerLeuAsnThrSerLeuLeuSerProProHisTyrGly 641
Db |||||
2383 TATCAGCAAGCTCGAAGCTTGGCGGATCCTAATACA----- 2421
Qy |||||
642 TyrGlnGlyLeuTrpThrProAsnTrpIleThrThrTyrThrIleThrLeuAsnAen 661
Db |||||
2421 ----- 2421
Qy -----
662 AsnSerSerAlaProThrSerAlaThrSerIleAlaGluGlnLysThrSerGluThr 681
Db -----
2421 ----- 2421
Qy -----
682 PheThrProSerAsnThrThrAlaSerIleProAsnIleLysAlaSerAlaGlySer 701
Db -----
2422 -----GCAAAATAATCGT-----CCTTATACTCTGAAGCT 2451
Qy |||||
702 GlySerGlySerAlaSerAsnSerGlyGluValThrIleThrLysHisThrLeuValVal 721
Db |||||
2452 ACATCGACTTAAACTGGGTATTAATCTGGCGCTGAGCGAGTAGCTTTTGGTTCCAAAT 2511
Qy |||||
722 AsnTrpAlaProValGlyTyriLeValAspProIleArgArgGlyAspLeuIleAlaAsn 741
Db |||||
2512 AGTTTATGGGATCCATTTTATGATATACGATCTGGCATTCAGCAATTCAGCAAGTGTG 2571
Qy |||||
742 SerLeu-----ValHisSer----- 746
Db -----
2572 GATGGCGCTCTTATTGTGAGGATTA-----TGGGTTCT 2607
Qy |||||
747 ---GlyArgAsnMetThrMetGlyLeuArgSerLeuLeuProAsnAsnSerTrpPheAla 765
Db |||||
2608 -----GGAGTTTCGAATTTCTTATCATGACCGGATGCTTTAGT---CAGGGA 2655
Qy |||||
766 LeuGlnGlyAlaAlaThrThrLeuPheThrLysGlnGlnLysArgLeuSerTyrHisGly 785
Db |||||
2656 TATCGTATATTAGTGGGTTATCTCTTAGGAGCAACTCTTACTTTGATCATCATG 2715
Qy |||||
786 TyrSerSerAlaSerLysGlyTyrThrValSerSerGlnAla-----SerGlyAla 802
Db |||||
2716 TTTGCTCTAGCATTTACCGAAGTATTGTTAGATCTTAAAGATTATGTTAGTGTCTCGTTC 2775
Qy |||||
803 HisGlyHisLysPheLeuLeuSerPheSerGlnSerSerAspLysMetLysGluLysGlu 822
Db |||||
2776 AATCATCATGCTTGATAGATCCGTTTATCTATCTATCTACCCACACAGCTTTTATGATCC 2835
Qy |||||
823 ThrAsnAsnArgLeuSerSerArgTyrTyrLeuSer-----AlaLeuCysPheGlu 839
Db |||||
2836 TATTTGTTCCGAGAT-----GCGTTTATC---CGTCTAGTACCGGTTTGGGAATCAG 2886
Qy |||||
840 HisProMetPheAspArgIleAlaLeuIleGlyAlaAlaAlaCysAsnTyrGlyThrHis 859
Db |||||
2887 CATATGAAACCTCATATACATTTGCGAGGAGGAGGATGTTGTTGGGATAATACTGT 2946
Qy |||||
860 AsnMetArgSerPheTyrGlyThrLysLysSerSerLysGlyLysPheHisSerThrThr 879
Db |||||
2947 CTGGCTGGAGATTTGCGCGGATACCGGATTTGATGATTACTCTCATCT---AAGCTCTAT 3003
Qy |||||
880 Leu-----GlyAlaSerLeuArgCysGluLeuArgAspSerMetProLeuArg 895
Db |||||
3004 TTGAATCAGTTCCTTTCTGCAAGCTGAGTTTCTTATGCGGATCATGATCTTTT 3063
Qy |||||
896 SerIleMetLeuThrProPheAlaGlnAlaLeuPheSerArgThrGluProAlaSerIle 915
Db |||||
3064 ACAGAGAGGCGCATCAGCTCCGGCATTC-----AAGACGCGCATCTCTTAAATCTA 3117
Qy |||||
916 ArgGluSerGlyAspLeuAlaArgPheThrLeuGluGlnAlaHisThrAlaValVal 935
Db |||||
3118 TCAGTTCTCTGAGTGAAGTTTGTATCGATGTTCTAGTACACATCCATAAT---AAATAT 3174
Qy |||||
936 Ser---ProIleGlyIleLysGlyAlaTyrSerSerAspThrTrpProThrLeuSerTrp 954
Db |||||
3175 AGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTGACTGAGACAG 3234
Qy |||||

```

Db      955 GluMetGluLeuAlaTyrGlnProThrLeuTyrTrpLysArgProLeuLeuAsnThrLeu 974
Qy      3235 CTCCTATCCCATCAAGAGACATGGACAACAGATGCTTTTCATTTTAGCAAGACATGAGATT 3294
Db      975 LeuIleGlnAsnAsnGlySerTrpValThrThrAsnThrProLeuAlaLysHisSerPhe 994
Qy      3295 GTGGTTAGAGATCTATGATGCTTCTCTAACAGTAATATAGAAGTATATGGCCATGGA 3354
Db      995 TyrGlyArgGlySer---HisSerLeuLysPheSerHisLeuLysLeuPheAlaAsnTyr 1013
Qy      3355 AGATGATGATGATCGAGATGCTTCTCGAGGCTATGCTTTCAGTGCAGGAGTAGAGTCCGG 3414
Db      1014 GlnAlaGluValAlaThrSerThrValSerHisTyrIleAsnAlaGlyGlyAlaLeuVal 1033
Qy      3415 TTC 3417
Db      1034 Phe 1034

RESULT 15
US-11-018-868-20
/ Sequence 20, Application US/11018868
/ Publication No. US20060034871A1
/ GENERAL INFORMATION:
/ APPLICANT: Chiron Corporation
/ APPLICANT: Grandi, Guido
/ APPLICANT: Ratti, Giulio
/ APPLICANT: Bonci, Alessandro
/ APPLICANT: Finco, Oretta
/ TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
/ FILE REFERENCE: 002441.00099 (P23152.001)
/ CURRENT APPLICATION NUMBER: US/11/018,868
/ CURRENT FILING DATE: 2004-12-22
/ NUMBER OF SEQ ID NOS: 186
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 20
/ LENGTH: 1034
/ TYPE: PRT
/ ORGANISM: Chlamydia trachomatis
US-11-018-868-20

Alignment Scores:
Pred. No.:      7, 41e-37      Length:      1034
Score:          522.50      Matches:      274
Percent Similarity: 37.1%      Conservative: 157
Best Local Similarity: 23.6%      Mismatches: 435
Query Match:      6.6%      Indels:      295
DB:              11      Gaps:      53

US-10-701-844-1 (1-4435) x US-11-018-868-20 (1-1034)
Qy      421 CTAGCTTATTTCTGCTGCTTTTAAATGGGGGGATATGCAGCAGAAATCATGTTCTCT 480
Db      7 LeuSerPheAlaCysLeuSerPhe-----PheTyrLeuSerThrIleSerIleLeu 23
Qy      481 CAGGAATTTACGATGGGGAGACGTTAACTGTATCATTT----- 519
Db      24 GlnAla-----AsnGluThrAspThrLeuGlnPheArgPheThrPheSerAsp 40
Qy      520 ---CCCTATCTGTTATAGGACATCCGAGTGGGACTACTGTTTCTTCGAGGAGTTA 576
Db      41 ArgGluIleGlnPheValLeuAspProAla-----SerLeuIleThrAlaGlnAsnIle 58
Qy      577 ACATTAATAAAATCTTGACAAATCTATTGCAGCTTTGCTTTCCTTTAAAGTTGTTGGGAACTTA 636
Db      59 ValLeuSerAsnLeuGlnSer-----AsnGly 67
Qy      637 TTAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTTCGAGACATACGGACT 696
Db      68 ThrGlyAlaCysThrIleSerGlyAsnThrGlnIlePheSerAsn----- 84
Qy      697 TCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGAAGTCTG-----TTTACT 747
Db      85 SerValAsnThrThrAlaAspSerGlyGlyAlaPheAspMetValThrThrSerPheThr 104

```

```

Qy      748 ATTGAGGGTTTAAAGAAATTATCTCTTTTCCAAATTCGAATTCATTACTTGCCTACTGCCT 807
Db      105 AlaSerAspAsnAlaAsnLeuPhe-----CysAsnAsnTyr----- 117
Qy      808 GCTCGAACGACTAATAAGGGTAGCCAGACTCCGACGACACACATCTACACCGCTCTAATGGT 867
Db      118 ---CysThrHisAsnLysGly-----GlyGly 125
Qy      868 ACTATTTATTTCTAAACACAGATCTTTTGTACTCAATATAGAGAAGTTCTCTATTCATAGT 927
Db      126 AlaIleArgSerGlyGlyProIleArgPheLeuAsnAsnGlnAspValLeuPheTyrAsn 145
Qy      928 AATTAGTCTCTGGA-----CATGGG 948
Db      146 AsnIleSerAlaGlyAlaLysTyrValGlyThrGlyAspHisAsnGluLysAsnArgGly 165
Qy      949 GGAGCTATAGATGCTTAAGAGCTTAACGCTTCAAGGAAATTAGCAAGCTTTGTGCTTCCAA 1008
Db      166 GlyAlaLeuTyrAlaThrThrIleThrLeuThrGly---AsnArgThrLeuAlaPheIle 184
Qy      1009 GAAAATATCTGCTCAAGCTGATGGGGAGCTTTGTCAAGTAGTCACCAGTTTCTCTGCT--- 1065
Db      185 AsnAsnMetSerGlyAspCysGlyGlyAlaIleSerAlaAspThrGlnIleSerIleThr 204
Qy      1066 -----ATGGCTAACGAGGCTCTTATTCGCC 1089
Db      205 AspThrValLysGlyIleLeuPheGluAsnAsnHisThrLeuAsnHisIleProTyrThr 224
Qy      1090 TTTGTAGCGAATGTTGCAGGAGTAGAGGGGGAGGATGCTGCTGTTTCAGATGGGGCAG 1149
Db      225 GlnAlaGluAsnMetAla-----ArgGlyGlyAlaIleCysSerArgArgAsp----- 240
Qy      1150 CAGGAGTGTCTATCTATCTTCAACAGAAAGTCCAGTAGTAAGTTTTCAGAGAAATACT 1209
Db      241 -----LeuCysSerIleSerAsnAsnSerGlyProIleVal----- 252
Qy      1210 CGGTAGATGTTGATGGGAACGTAGCCGAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1269
Db      253 -----PheAsnTyrAsnGlnGlyGlyLysGlyGlyAlaIleSerAlaThrArgCys 269
Qy      1270 GTTGCTTTCTCGAATAATGGAACACCTTTGTTCTCAACAATGTTGCTTCTCTCTGTTTAC 1329
Db      270 ValIleAspAsnAsnLysGluArgIleIlePheSerAsnAsn----- 283
Qy      1330 ATTGCTGCTAAGCAACCAAGTGCAGACAGCTTCTAATACGAGTAATAATTAACGAGAT 1389
Db      284 -----SerSerLeuGlyTrpSerGlnSerSerSerAlaSerAsn 296
Qy      1390 GGAGGAGCTATCTCTGTGAAGATGGTCGCGACAGCAGGATCCAATACTCTGGATCAGTT 1449
Db      297 GlyGlyAlaIleGlnThrThrGlnGlyPheThrLeu---ArgAsnAsnLysGlySerIle 315
Qy      1450 TCCTTTGATGAGGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGA 1509
Db      316 TyrPheAsp-----SerAsnThrAlaThrHisAlaGlyGly 327
Qy      1510 GCTATTTATGCCAAAAAGCTCTCGGTTGCTTAACCTGTGCTGCTGTCACAAATTTTAAAGGAT 1569
Db      328 AlaIleAsnCysGlyTyrIleAspIleArgAspAsnGlyProValTyrPheLeuAsnAsn 347
Qy      1570 ATCGCTAATGATGGTGGAGCGAATTTATTAGAGAA-----TCTGGAGAGCTCAGTTTA 1623
Db      348 SerAlaIleTyrGlyAlaAlaPheAsnLeuSerLysProArgSerAlaThrAsnTyrIle 367
Qy      1624 TCTGCTGATTATGAGATATTATTTC-----GATGGGAT 1659
Db      368 HisThrGlyThrGlyAspIleValPheAsnAsnAsnValValPheThrLeuAspGlyAsn 387
Qy      1660 CTTAAAGAACACCAAGAGAGATGCTGCCGATGTTAATGGCGTAACCTGTGCTCCTCACAA 1719
Db      388 Leu-----LeuGlyLysArgLysLeuPheHisIleAsnAsnAsnGluIleThrProTyr 405

```



```

1987 CAGACAGGTGGAGT---CTGTATATGGAAGCTGGAGTACATGGATTTGTAACTCCA 2043
1240 GlnLysGluGlySerLysLeuMetLysProGlyAlaValLeu-----1254
2044 CAACACACACAGAGCTCTGCGCTTAATCATGATGATCAGCTTTCCAACTCTGCAATTTG 2103
1255 -----SerAsnGlnAsnLeuAlaAsnGlyAlaLeuValLeuAsnGlyLeuThrIle 1271
2104 TCTCTTTCTTTGTTAGCAACAATGCA-----GTTACGAATCCT-----2145
1272 AspLeuSerSerMetGlyThrProGlnAlaGlyLeuPheSerProGluLeuArg 1291
2146 -----CCTACC 2151
1292 IleValAlaThrThrSerSerAlaSerGlyGlySerGlyValSerSerSerIleProThr 1311
2152 AATCCT-----CCAGCGCAAGATTCTCATCTCCGATGATGATGATGATGATGATGATGAT 2196
1312 AsnProLysArgIleSerAlaAlaAlaProSerGlySerAlaAlaThrThrProThrMet 1331
2197 ACTGCTGGTCTGTACAAATAGTGGCTATCTTTTGGAGATTGGATGATGATGATGATGATGAT 2256
1332 SerGluAsnLysValPheLeuThrGlyAspLeuThrLeuIleAspProAsnGlyAsnPhe 1351
2257 TATGATAGTATGATGCTAGTCTTCTAATCAAAAATCAATGCTCTGAAATACAGTTA 2316
1352 TyrGlnAsn---ProMetLeuGlySerAspLeuAspValProLeuIleLysLeu-----1368
2317 GGGACTAAGCCCGCAGCTAATGCCCATCA-----GATTGACTCTA-----GGG 2361
1369 -----ProThrAsnThrSerAspValGlnValThrAspLeuThrLeuSerGly 1384
2362 AATGAGATGCTAATGATGCTATCAAGGAGCTGGAAGCTGCGGGATCTTAATACA 2421
1385 AspLeuPheProGlnLysGlyThrMetGlyThrTrpThrLeuAspSerAsnProGlnThr 1404
2422 GCMAATAGTCTTATCTACTCTGAAGAGTACATGAGCTAATAAACTGGGTATATCTCT--- 2478
1405 GlyLys-----LeuGlnAlaArgTrpThrPheAspThrThrArgArgTrp 1419
2479 -----GGGCTCGAGCGAGTA 2493
1420 ValTyrIleProArgAspAsnHisPheTyrAlaAsnSerIleLeuGlySerGlnAsnSer 1439
2494 GCTTCTTTGTTCCAAATAGTTTATGGGATCCATTTTAGATATACGATCTGCGCATCA 2553
1440 MetIleValValLysGlnGlyLeuIleAsnAsnMetLeuAsn-----1453
2554 GCMAATCAAGCAAGTGTGGATGGCGCTCTTATGTCGAGGATATGGTTCTGGAGTT 2613
1454 -----AsnAlaArgPheAspAspIleAlaTyr---AsnAsnPheTrpValSerGlyVal 1470
2614 TCGAATTTCTCTCATGACCGGATGCTTTAGTCCAGGATATCGTATATAGTGGG 2673
1471 GlyThrPheLeuAlaGlnGlnGlyThrProLeuSerGluGluPheSerTyrTyrSerArg 1490
2674 GGTATTTCTCTTAGGACAAATCTCTACTTTGGATCATCG---ATGTTTGGCTTAGCATTTT 2730
1491 GlyThrSerValAlaIleAspAlaLysProArgGlnAspPheIleLeuGlyAlaAlaPhe 1510
2731 ACCGAAGTATTTGGTAGATCAAGATTATGTAGTGTGCTTCCCAATCATCATGCTTGC 2790
1511 SerLysMetValGlyLysThrLys-----AlaIleLysLysMetHisAsnTyrPheHis 1528
2791 ATAGGATCCGTTTATCTATCTACCAACAGCTTTATGTGATCTTATTTGTCGGAGAT 2850
1529 LysGlySerGluTyrSerTyrGlnAlaSerValTyrGlyLysPheLeuTyr-----1546
2851 GCGTTTATCCGT-----GCT 2865
1547 ---PheLeuLeuAsnLysGlnHisGlyTrpAlaLeuProPheLeuIleGlnGlyValVal 1565

2866 AGCTACGGGTTTGGGAATCAGCATATGAAACCTCATAT---ACATTTTCAGAGAGAGC 2922
1566 SerTyrGlyHisIleLysHisAspThrThrThrLeuTyrProSerIleHisGluArgAsn 1585
2923 GATGTCCTTGGGATAATAACTGCTGCTGGAGATGGGCGGATACCGATTTG 2982
1586 LysGlyAspTrpGluAspLeuGlyTrpLeuAlaAspLeuArgIleSerMetAspLeuLys 1605
2983 ATTACTCCTCATTAAGCTCTATTGTAATGAGTTGCTCTCTCGTTCGCTCAAGCTGAGTTTCT 3042
1606 ---GluProSerLysAspSerSerLysArgIleThrValTyrGlyLeuGluTyrSer 1624
3043 TATGCCGATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCGCATTTCAAGACGGA 3102
1625 SerIleArgGlnLysGlnPheThrGluIleAspTyrAspProArgHisPheAspAspCys 1644
3103 CATCTCCTAATCTATCATGTTCTCTGCTGGAGTGAAGTTTGAAT-----CGATCT 3150
1645 AlatyArgAsnLeuSerLeuProValGlyCysAlaValGluGlyAlaIleMetAsnCys 1664
3151 TCTAGTACACATCCTAATAATATAGCTTTATGGGCGCTTATATCTGTGATGCTTATCGC 3210
1665 AsnIleLeuMetTyrAsnLysLeuAlaLeu-----AlatyMetProSerIleTyrArg 1682
3211 ACCATCTCTGCTACTGACAGAACGCTCTATCCCATCAAGACATCGGACACAGATGCC 3270
1683 AsnAsnProValCysLysTyrArgValLeuSerSerAsnGluAlaGlyGlnValIleCys 1702
3271 TTTCAATTAGCAAGACATGGAGTGTGTGTAGAGATCTATGATGCTTCTCTTCAACAGT 3330
1703 -----GlyValProThrArgThrSerAlaArgAlaGluTyrSerThr 1716
3331 AATATA-----GAAGTATATGGCAT-----GGAAGA 3357
1717 GlnLeuTyrLeuGlyProPheThrThrLeuTyrGlyAsnTyrThrIleAspValGlyMet 1736
3358 TATGATGATCGAGATGCTTCTCGAGGCTATGTTGAGTGCAGGAAGTAGTCCGGTTC 3417
1737 TyrThrLeuSerGlnMetThr-----SerCysGlyAlaArgMetIlePhe 1751

RESULT 17
US-11-018-868-16
; Sequence 16, Application US/11018868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Pinco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 1751
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-16

Alignment Scores:
Pred. No.: 1.19e-32 Length: 1751
Score: 475.50 Matches: 269
Percent Similarity: 36.6% Conservative: 148
Best Local Similarity: 23.6% Mismatches: 454
Query Match: 6.0% Indels: 269
DB: 11 Gaps: 51

US-10-701-844-1 (1-4435) x US-11-018-868-16 (1-1751)
QY 484 GGAATTTACGATGGGAGAGCTTAACGTATCATTTCCCTATCTGTTATAGGAGATCCG 543
```

Db	719	GlyIleTyrAlaLysLysAlaLysMetSerArgIleAspGlnLeuAsnIleSerGluAen	738
Qy	544	AGTGGGACTACTGTTTTTTCGCGAGGAGAGTTAAACATTAAAAAAATCTTGACAATTTCTATT	603
Db	739	SerAlaThrGluIle-----GlyGlyGlyLeuValGlyLys-----GluGlyGly	756
Qy	604	GCAGCTTTGCTTTAAAGTCTTTTGGAACTTATTAGGGAGTTTACTGTTTGGGAGAGA	663
Db	757	AlaLeuValSerLeuSerValThrGluAsnLeuValGlyLys-----GluGlyGly	773
Qy	664	GGACACTCGTTGACTTTCGAGAACATACAGGACTTCTACAAATGGGCGAGCTCTAAGTAAT	723
Db	774	GlyLeuHisAlaLysThrValAsnIleSerAsnLeuLysSerGlyPheSerPheSerAsn	793
Qy	724	AGCGTCTGATGGACTGTTTACTATTGAGGGTTTAAAGAAATATCTTTTCCCAATTCG	783
Db	794	AsnLysAlaAsnSerSerThr-----GlyValAlaThrThr	806
Qy	784	AAATCATTACTTGCCTGACTCCCTGCTCAACGACTAATAAGGTAGCCAGACTCCGACG	843
Db	807	AlaSerAlaProAlaAlaAlaAlaSerLeuGlnAlaAlaAlaValProSer	826
Qy	844	ACAACATCTACACGCTTAATGGTACTATTATTCTTAAACACAGCTTTTGGTTACTCAAT	903
Db	827	SerProAlaThrProThr-----TyrSerGlyVal-----GlyGlyAlaIleTyrGly	832
Qy	904	AATGAGAAGTCTCATTTCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCT	963
Db	833	-----TyrSerGlyVal-----GlyGlyAlaIleTyrGly	843
Qy	964	AAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTCTCAAGAAAATACTGTCT	1020
Db	844	GluLysValThrPheSerGlnCysSerGlyThrCysGlnPheSerGlyAsnGlnAlaIle	863
Qy	1021	-----CAAGTGAATGGGGAGCTTGTCAAGTAGTCACC	1053
Db	864	AspAsnAsnProSerGlnSerSerLeuAsnValGlnGlyAlaIleTyrAlaLysThr	883
Qy	1054	AGTTTCTCTGCTATGGCTTAACGAGGCTCTATTGCTTCTGTA-----	1095
Db	884	SerLeuSerIleGlySerSerAspAlaGlyThrSerTyrIlePheSerGlyAsnSerVal	903
Qy	1096	-----GCCAATGTTTCGAGGAGTAGAGGGGAGGAGATTGCTGCTGTTCCAGGAT	1143
Db	904	SerThrGlyLysSerGlnThrThrGlyGlnIleAlaGlyLys-----	917
Qy	1144	GGGAGAGGGAGGTGCATCATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGA	1203
Db	918	-----AlaIleTyrSerProThrValThrLeuAsnCysProAlaThrPheSerAsn	934
Qy	1204	AATACTCG-----GTAGAGTTGATGGGAACGTAGCCCGA-----	1239
Db	935	AsnThrAlaSerMetAlaThrProLysThrSerSerGluAspGlySerSerGlyAsnSer	954
Qy	1240	-----GTAGGAGGAGGAGATTACTCTCTACGGAACGTGCTTCTCTCGAATAAT	1287
Db	955	IleLysAspThrIleGlyGlyAlaIle-----AlaGlyThrAlaIleThrLeuSer	971
Qy	1288	GGAAAAACCTTTGTTCTCAACAATGTTGCT-----TCTCCTGTTTACATGCTGCT	1338
Db	972	GlyValSerAspPheSerGlyAsnThrAlaAspLeuGlyAlaAlaIleGlyThrLeuAla	991
Qy	1339	AAGCAACCAACAAGTGGACGGCTCTTAATACGAGTAATAATTACGGAGAT-----	1389
Db	992	AsnAlaAsnThrProSerAlaThrSerGlySerGlnAsnSerIleThrGluLysIleThr	1011
Qy	1390	-----GGAGGAGCTATCTTCTGTAAAGAAATGGTGG-----CAAGCAGGATCCAAATAAC	1437
Db	1012	LeuGluAsnGlySerPheIlePheGluArgGlnAlaAsnLysArgGlyAlaIleTyr	1031
Qy	1438	TCTGGATCAGTTTCTTGTATGGAGGAGGAGTAGTTTCTTTTATGAGCAATGTAGCTGCT	1497

Db	1500	ThrSerGluGluPheThrTyrTyrSerArgGlyAlaSerValAlaLeuAspAlaLysPro	1519	
Qy	2704	GGATCATCGATGTTT---GGTCTAGCATTTACCGAAGTATTGCTAGATCTAAAGATTAT	2760	
Db	1520	AlaHisAspValIleValGlyAlaAlaPheSerLysMetIleGlyLysThrLysSer---	1538	
Qy	2761	GTAGTGTGTCGTTCCAAATCATCATGCTTGCATAGATCGTTTATCTATCTATCCCAACAA	2820	
Db	1539	---LeuLysArgGluAsnAsnTyrThrHisLysGlySerGluTyr---SerTyrGlnAla	1556	
Qy	2821	GCTTATGTGGATCCCTATTGTTGCGAGATGCGGTTTATCCGT-----	2862	
Db	1557	SerValTyrGlyGlyLysProPheHisPheValIleAsnLysThrGluLysSerLeu	1576	
Qy	2863	-----GCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCA	2901	
Db	1577	ProLeuLeuGlnGlyValIleSerTyrGlyTyrIleLysHisAspThrValThrHis	1596	
Qy	2902	TAT---ACATTTGCGAGGAGCGATGTTGTTGGATAATAACTGCTGCGTGGAGAG	2958	
Db	1597	TyrProThrIleArgGluArgAsnLysGlyGluTrpGluAsp-----LeuGlyTrp	1613	
Qy	2959	ATTGGAGCG-----GGATTACCGATTGTGATTACTCCATTAAGCTCTATTGAATGAG	3012	
Db	1614	LeuThrAlaLeuArgValSerSerValLeuArgThrProAlaGlnGlyAspThrLysArg	1633	
Qy	3013	TTGCGTCTTTCGTCGCAAGCTAGTTCCTTATGCGCATCATGAATCTTTTACAGAGGAA	3072	
Db	1634	IleThrValTyrGlyGluLeuGluTyrSerSerIleArgGlnLysGlnPheThrGluThr	1653	
Qy	3073	GCGCATCAAGCTCGGCATTCAAGAGCGGCACATCTCTAAATCTATCAGTTCCTGTGGA	3132	
Db	1654	GluTyrAspProArgTyrPheAspAsnCysThrTyrArgAsnLeuAlaIleProMetGly	1673	
Qy	3133	GTCAAGCTTTGAT---CGATGTTCTCTAGTACACATCTCTAAATAATATAGC---TTTATGGCG	3186	
Db	1674	LeuAlaPheGluGlyGluLeuSerGlyAsnAspIleLeuMetTyrAsnArgPheSerVal	1693	
Qy	3187	GCTTATATCTGTGATGCTTATCCGACCATCTCTGCTAGTACGACACAGCTCTCCAT	3246	
Db	1694	AlaTyrMetLeuSerIleTyrArgAsnSerProThrCysLysTyrGlnValLeuSerSer	1713	
Qy	3247	CAAGACATGGACACAGATGCTTCTATTAGCAAGATCGAGTGTGGTTAGAGGA	3306	
Db	1714	GlyGluGlyGlyGluIleLeuCys-----GlyValProThrArgAsn	1727	
Qy	3307	TCTATGATGCTTCTCTAAACAGTAATATA-----GAAGTATATGGC	3348	
Db	1728	SerAlaArgGlyGluTyrSerThrGlnLeuTyrLeuGlyProLeuTyrThrLeuTyrGly	1747	
Qy	3349	CATGGAAGATATCATGATATCGAGATGCTTCTCGAGGCTATGGTTGAGTCGAGAGTAGA	3408	
Db	1748	SerTyrThrIleGluAlaAspAlaHisThrLeuAlaHisMetMetAsnCysGlyAlaArg	1767	
Qy	3409	GTCCGGTTC	3417	
Db	1768	MetThrPhe	1770	
RESULT 19				
US-11-018-868-17				
; Sequence 17, Application US/11018868				
; Publication No. US20060034871A1				
; GENERAL INFORMATION:				
; APPLICANT: Chiron Corporation				
; APPLICANT: Grandi, Guido				
; APPLICANT: Ratti, Giulio				
; APPLICANT: Bonci, Alessandro				
; APPLICANT: Finco, Oretta				
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis				
; FILE REFERENCE: 002441.00099 (pp23152.001)				
; CURRENT APPLICATION NUMBER: US/11/018,868				
; CURRENT FILING DATE: 2004-12-22				
; NUMBER OF SEQ ID NOS: 186				

; SOFTWARE: PatentIn version 3.2				
; SEQ ID NO 17				
; LENGTH: 1770				
; TYPE: PRT				
; ORGANISM: Chlamydia trachomatis				
US-11-018-868-17				
Alignment Scores:				
Pred. No.:	4,19e-27	Length:	1770	
Score:	412.00	Matches:	258	
Percent Similarity:	35.5%	Conservative:	148	
Best Local Similarity:	22.6%	Mismatches:	421	
Query Match:	5.2%	Indels:	316	
DB:	11	Gaps:	51	
US-10-701-844-1 (1-4435) x US-11-018-868-17 (1-1770)				
Qy	673	TTGACTTTCGAGAACATACGGACTTCTACA-----	708	-----AATGGG
Db	716	IleThrAspGluSerValSerSerSerGluSerGlySerThrProGlnAspGly	735	-----
Qy	709	GCAGCTTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTAAAGAAATTA	768	-----
Db	736	GlyAlaAlaSerSerGlyAlaProSerGlyAspGlnSerIle-----	749	-----
Qy	769	TCCTTTTCCAAATTGCAATTCAATTACTTGCCGTACTGCTGCTGCAACGACT-----	819	-----
Db	750	-----SerAlaAsnAlaCysLeuAlaLysSerTyrAlaAlaSerThrAspSerSer	766	-----
Qy	820	-----AATAAGGTAGCCAGACTCCGACGACAACTATACACCGTCTAATGCT	867	-----
Db	767	ProValSerAsnSerSerGlySerGluProValThrSerSerSerAspSerAspVal	786	-----
Qy	868	ACTATTATTCTTAAACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTTCTATAGT	927	-----
Db	787	ThrAlaSerSerAsp-----AsnProAspSerSerSer-----	797	-----
Qy	928	AATTGATGCTCTGGAGAT-----	945	-----
Db	798	-----SerGlyAspSerAlaGlyAspSerGluProThrGluProGluAlaGly	814	-----
Qy	946	-----GGGGAGCTATAGATGCTAAGAGCTTAACG	975	-----
Db	815	SerThrThrGluThrLeuThrLeuIleGlyGlyAlaIleTyrGlyGluThrValLys	834	-----
Qy	976	GTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAAATACTGCT-----CAAGCT	1026	-----
Db	835	IleGluAsnPheSerGlyGlnGlyIlePheSerGlyAsnLysAlaIleAspAsnThrThr	854	-----
Qy	1027	GATGGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGCTATGGCTAACGAGGCTCCTATT	1086	-----
Db	855	GluGlySerSerLys-----	860	-----
Qy	1087	GCCTTTGTAGCGAATGTTTCAGGAGTAAGAGGGGAGGAGTCTGCTGTTAGGATGGG	1146	-----
Db	861	-----SerAspValLeuGlyAlaValTyrAlaLysThrLeuPhe	874	-----
Qy	1147	CAGCAGGGAGTGCATCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTCCAGAAAT	1206	-----
Db	875	AsnLeuAspSerGlySerSerArgArgThr-----ValThrPheSerGlyAsn	890	-----
Qy	1207	ACT-----GCGGTAGAGTTTGATGGACGTAGCCGAGTAGGAGGAGGATTTACTCC	1260	-----
Db	891	ThrValSerSerGlnSerThrThrGlyGlnValAla-----GlyGlyAlaIleTyrSer	908	-----
Qy	1261	-----TACCGGAACTGCTTCTCTGAATAATGAAACCTGTTTCTC	1305	-----
Db	909	ProThrValThrIleAlaThrProValValPheSerLysAsnSerAlaThr-----	925	-----
Qy	1306	AACAATGTTGCTTCTCTGTTTACATTTGCTGCTAAGCAACCAAGTGCACAGGCTTCT	1365	-----
Db	926	AsnAsnAlaAsnAsnThrThrAspThrGlnArgLysAspThrPheGlyGlyAlaIleGly	945	-----

QY 1366 AATACGAGTAATATTACGGAGATGGAGGCTATCTTCTGTAGAAGTGGCGCAAGCA 1425
Db |||||
946 AlaThrSerAlaValSerLeuSerGlyGlyAlaHisPheLeuGluAsnValAlaAspLeu 965
QY 1426 GGATCC-----AATAACTCTGGATCGATGTTCTTCTTGATGGA 1461
Db |||||
966 GlySerAlaIleGlyLeuValProGlyThrGlnAsnThrGluThrValLysLeuGlu--- 984
QY 1462 GAGGAGTAGTTTCTTTAGTAGCAATAGTCTCTGGAAAGGGGAGCTATTATATGCC 1521
Db |||||
985 SerGlySerTyThrPheGluLysAsnLysAlaLeu---LysArgAlaThrIleTyrAla 1003
QY 1522 AAAAGCTCTGGTTCCTAAGCTGGCTGCTGACAAATTTTAAAGCAATATCGCT---AAT 1578
Db |||||
1004 ProValValSerIle---LysAlaTyrThrAlaThrPheAsnGlnAsnArgSerLeuGlu 1022
QY 1579 GATGTGGAGCGATTTATTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTTATGGA 1638
Db |||||
1023 GluGlySerAlaIleTyrPheThrLysGluAlaSerIleGlu-----SerLeuGly 1039
QY 1639 GATATTATTTCGATGGGAATCTT-----AAAAGAACCA 1671
Db |||||
1040 SerValLeuPheThrGlyAsnLeuValThrLeuThrLeuSerThrThrThrGluGlyThr 1059
QY 1672 GCCAAGAGAAATCGCCGATGTTAAT-----GGCGTAACGTGTCTCTCAAGCCATT 1725
Db |||||
1060 ProAlaThrThrSerGlyAspValThrLysTyrGlyAlaAlaIlePheGlyGlnIleAla 1079
QY 1726 TCG-----ATCGGATCGGAGGGAAA 1746
Db |||||
1080 SerSerAsnGlySerGlnThrAspAsnLeuProLeuLysLeuIleAlaSerGlyGlyAsn 1099
QY 1747 ATA----- 1749
Db |||||
1100 IleCysPheArgAsnAsnGluTyrArgProThrSerSerAspThrSerGlyThrSerThrPhe 1119
QY 1750 -----ACGACATTAAGCTAAAGCAGGAGGATCAGATTCTC 1785
Db |||||
1120 CysSerIleAlaGlyAspValLysLeuThrMetGlnAlaAlaLysGlyLysThrIleSer 1139
QY 1786 TTTAATGATCCATCCAGATGCA---AACGGAAATAACCGCAGCGAGTCTCCAAA 1842
Db |||||
1140 PhePheAspAlaIleArgThrSerThrLysLysThrGlyThrGlnAlaThrAlaTyrAsp 1159
QY 1843 CTTCTAAAAATPAACGATGGTGAA-----GGATACACAGGGAT 1881
Db |||||
1160 ThrLeuAspIleAsnLysSerGluAspSerGluThrValAsnSerAlaPheThrGlyThr 1179
QY 1882 ATTGTTTTTGCTAAT-----GGAACGACTTTGTATCCAAAATGTTACGATA 1929
Db |||||
1180 IleLeuPheSerSerGluLeuHisGluAsnLysSerTyIleProGlnAsnValValLeu 1199
QY 1930 GAGCAAGGAGGATTTCTTCGTGAAAAGGCAAAATATCATGTAATCTCTAAAGTCAG 1989
Db |||||
1200 HisSerGlySerLeuValLeuLysProAsnThrGluLeuHisValIleSerPheGluGln 1219
QY 1990 ----ACAGGTGGGAGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCAAA 2046
Db |||||
1220 LysGluGlySerSerLeuValMetThrProGlySerValLeuSerAsnGlnThr----- 1237
QY 2047 CCACCAACACGCTCTCCGCTAATCATGTTGATCACGCTTTCCATCTGATTTGTCT 2106
Db |||||
1238 -----ValAlaAspGlyAlaLeuValIleAsnAsnMetThrIleAsp 1251
QY 2107 CTTTCTCTTTTG-----TTAGCAAAACAATGCAAGTTAGCAATCTCTCTACC--- 2151
Db |||||
1252 LeuSerSerValGluLysAsnGlyIleAlaGluGlyAsnIlePheThrProGluLeu 1271
QY 2152 -----AATCTCCAGCGCAAGATTCTCAT--- 2175
Db |||||
1272 ArgIleIleAspThrThrThrGlyGlySerGlyGlyThrProSerThrAspSerGluSer 1291
QY 2175 ----- 2175

Db 1292 AsnGlnAsnSerAspAspThrGluGluGlnAsnAsnAspAlaSerAsnGlnGlyGlu 1311
QY 2176 -----CTGCAGTCTCAATCGT----- 2190
Db |||||
1312 SerAlaAsnGlySerSerProAlaValAlaAlaAlaHisThrSerArgThrArgAsn 1331
QY 2191 -----AGCACACTGCT 2202
Db |||||
1332 PheAlaAlaAlaAlaThrAlaThrProThrThrThrProThrAlaThrThrThrSer 1351
QY 2203 GGTTCGTTCACAAATTAGTGGCCCTATCTTTTTCAGGATTTGGATGACACAGCTATTATCAT 2262
Db |||||
1352 AsnGlnValIleLeuGlyGlyGlyLeuLeuLeuAspProAsnGlyThrPhePheGln 1371
QY 2263 AGGTATGATGGCTAGGTTCTTAATCAAAAATCAATGTCCTGAATAATTACAGTTAGGACT 2322
Db |||||
1372 Asn---ProAlaLeuArgSerAspGlnIleSerLeuLeuValLeuProThrAspSer 1390
QY 2323 AAGCCCCAGCTAATGCCCATCAGATTGTGACTCTAGGGAATGAGATGCCATAAGTATGCC 2382
Db |||||
1391 Ser---LysMetGlnAlaGlnLysIleValLeuThrGlyAspIleAlaProGlnLysGly 1409
QY 2383 TATCAAGGAAGCTCGAAGCTTGGTGGATCCTTAATACAGCAAAATAAGTGCCTTATATCAT 2442
Db |||||
1410 TyrThrGlyThr-----LeuThrLeuAspProAspGlnLeuGlnAsnGly-----Thr 1425
QY 2443 CTGAAGCTACATGACTAAAACTGGGTATATCTCTGGCCCTGAGCGAGTACTTCTTTG 2502
Db |||||
1426 IleSerValLeuTyrLysPheAspSerTyr-----ArgGlnTyrAlaTyr 1440
QY 2503 GTTCCA-----AATAGTTTATGGGGATCCATTTTAGATATACGA 2541
Db |||||
1441 ValProArgAspAsnHisPheTyrAlaAsnSerIleLeuGlySerGlnMetLeuMetVal 1460
QY 2542 TCTGGCATTCAGCAATT-----CAAGCAAGTGTGGATGGGGCTCT 2583
Db |||||
1461 ThrValLysGlnGlyLeuLeuAsnAspLysMetAsnLeuAlaArgPheGluGluValSer 1480
QY 2584 TATTGTCGAGGATATGGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCGCATGCT 2643
Db |||||
1481 Tyr---AsnAsnLeuTyrIleSerGlyLeuGlyThrMetLeuSerGlnValGlyThrPro 1499
QY 2644 TAGTGTACGATATCGTATATATTAGTGGGGTATTCTTCTTAGGAGCAACTCCTACTTT 2703
Db |||||
1500 ThrSerGluGluPheThrTyrTyrSerArgGlyAlaSerValAlaLeuAspAlaLysPro 1519
QY 2704 GGATCATCGATGTTT---GGTCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTAT 2760
Db |||||
1520 AlaHisAspValIleValGlyAlaAlaPheSerLysMetIleGlyLysThrLysSer--- 1538
QY 2761 GTAGTGTGCTGCTCAATCATCTTGCATAGATCCGTTTATCTATCTACCCACAA 2820
Db |||||
1539 ---LeuLysArgGluAsnAsnTyrThrHisLysGlySerGlyTyr---SerTyrGlnAla 1556
QY 2821 GCTTTATGTGATCTCTATTGTTGCGAGATGGGTTTATCCGT----- 2862
Db |||||
1557 SerValTyrGlyGlyLysProPheHisPheValIleAsnLysLysThrGluLysSerLeu 1576
QY 2863 -----GCTAGCTACGGTTTGGGAATCAGCATATGAAAACCTCA 2901
Db |||||
1577 ProLeuLeuGlnGlyValIleSerTyrGlyTyrIleLysHisAspThrValThrHis 1596
QY 2902 TAT---ACATTTTCAGAGAGAGCGATGTTCTGGGATAATAACTGTCTGGCTGGAGAG 2958
Db |||||
1597 TyrProThrIleArgGluAsnLysGlyGluTyrGluAsp-----LeuGlyTyr 1613
QY 2959 ATTGAGGAG-----GGATTACGATTGTGATTCTCATCTTAAGCTCTATTGTGATGAG 3012
Db |||||
1614 LeuThrAlaLeuArgValSerSerValLeuArgThrProAlaGlnGlyAspThrLysArg 1633
QY 3013 TTGCGTCTCTTTCGTCAAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAA 3072
Db |||||

```
Db      1634  IleThrValTyrGlyGluLeuGluTyrSerSerIleArgGlnLysGlnPheThrGluThr 1653
QY      3073  GCGGATCAAGCTCGGCATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTGTGGA 3132
Db      1654  GluTyrAspProArgTyrPheAspAenCysThrTyrArgAsnLeuAlaIleProMetGly 1673
QY      3133  GTGAAGTTTGAT---CGATGTTCTAGTACACATCTCTAAATAATATAGC---TTTATGGCG 3186
Db      1674  LeuAlaPheGluGlyGluLeuSerGlyAenAspIleLeuMetTyrAsnArgPheSerVal 1693
QY      3187  GCTTATATCTGTGATGCTTATCGCACCATCTCTGTGACTGAGACAAAGCTCTATCCCAT 3246
Db      1694  AlaTyrMetLeuSerIleTyrArgAenSerProThrCysLysTyrGlnValLeuSerSer 1713
QY      3247  CAAGAGACATGGAACAACATGCCITTCATTAGCAAGACATGGAGTGTGTTAGAGCA 3306
Db      1714  GlyGluGlyGlyGluIleCys-----GlyValProThrArgAen 1727
QY      3307  TCTATGTATGCTTCTCAACAGTAATATA-----GAAGTATATGGC 3348
Db      1728  SerAlaArgGlyGluTyrSerThrGlnLeuTyrLeuGlyProLeuTrpThrLeuTyrGly 1747
QY      3349  CATGGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGAGTCGAGGAAGTAGA 3408
Db      1748  SerTyrThrIleGluAlaAspAlaHisThrLeuAlaHisMetAenCysGlyAlaArg 1767
QY      3409  GTCCGGTTC 3417
Db      1768  MetThrPhe 1770
```

RESULT 20

```
US-11-052-554A-1
; Sequence 1, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1268
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-1
```

```
Alignment Scores:
Pred. No.:      4,43e-15      Length:      1268
Score:          273.50      Matches:      304
Percent Similarity: 33.2%      Conservative: 187
Best Local Similarity: 20.6%      Mismatches: 495
Query Match:      3.5%      Indels:      493
DB:              11      Gaps:      79
```

US-10-701-844-1 (1-4435) x US-11-052-554A-1 (1-1268)

```
QY      415  ATGATTTAGTCTATTCTGCTCTTTAAATGGGGGGATGATGAGCGAGAAATCATG 474
Db      21  MetMetLeuSerPheProValAlaSer-----GlnAlaAlaGlyLeuVal 35
QY      475  GTTCCTCAAGA---ATTACGATGGGAGACGGTTAACTGATCATCTTCCCTATCTGTT 531
Db      36  IleLysAenGlyThrValTyrAsnAlaAen-----GlyValProValValAsp 51
QY      532  ATAGGAGATCGAGTGGGACTACTGTTTCTGCGAGGAGGTAAACATTAATAAAATCTT 591
```

```
Db      52  IleAenLysProAenGlySerGlyLeu---SerHisAenIleTrpAspAenLeuAenVal 70
QY      592  GAC-----AATTCATTATGAGCTTTTGCCTTTTAAAGTTGTTTGGG 630
Db      71  AspLysAenGlyValValPheAenAenSerAlaAenGluSerSerThrSerLeuAlaGly 90
QY      631  AACTTATTAGGAGTGTATGTTTATGGGAGAGGACACTCGTGACTTTCGAGAACATA 690
Db      91  AenIleGlnGlyAen---SerAenLeuThrSerGlySerAlaLysValIleLeuAenGlu 109
QY      691  CGGACTTCTACAAATGGGCGACGCTCTAAGTAATAGCGTGTGATGGAGCTGTTTACTATT 750
Db      110  ValThrSerLysAenProSer-----ThrIle 118
QY      751  GAGGCTTTTAAAGAAATTATCTTTTCCAATTCCAATTTCATTACTTGCCTACTGCTGCT 810
Db      119  AenGlyMetMetGluValAlaGlyAspLysAlaAspLeuIleAlaAenProAenGly 138
QY      811  GCAACGACTTAATAGGCTAGCCAGACTCGAGC-----ACAACATCTACA 855
Db      139  IleThrValAenGlyGlySerIleAenThrGlyLysLeuThrLeuThrThrGlyThr 158
QY      856  CCG-----TCTAATGGTACTATT----- 873
Db      159  ProAspIleGlnAspAspLysLeuAlaGlyTyrSerValAenGlyGlyThrIleThrLeu 178
QY      874  -----TATCTAAACAGATCTTTTG-----TTACTCAATAAT 906
Db      179  GlyLysLeuAspAenAlaSerProThrGluIleLeuSerArgAenValValValAenGly 198
QY      907  GAGAAGTTCCTCATTTCTATAGTAATTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAG 966
Db      199  LysValSerAlaAspGluLeuAenValValAlaGlyAen---AsnTyrValAenAlaAla 217
QY      967  AGCTTAACGTTCAAGCAATTAGCAAGCTTGTGCTTCCACAGAAATAGTCTCAAGCT 1026
Db      218  GlyGlnValThrGlySerValSerAla-----ThrGlySerArg 230
QY      1027  GATGGGGAGCTGTGCAAGTAGTCACACAGTTCCTCTGCTATGCTACGACGAGCTCTATT 1086
Db      231  AsnGlyTyrSerValAspValAlaLysLeuGlyGlyMetTyrAlaAenLysIleSerLeu 250
QY      1087  GCCTTTGTAGCGAATTTGCGAGGATPACAGGGGAGGAGGATGCTCTGTTCAAGATGGG 1146
Db      251  ValSerThrGluLysGlyValGlyValArgAenLeuGlyValIleAla-----GlyGly 268
QY      1147  CAGCAGGAGTGTCA----- 1161
Db      269  ValAenGlyValSerIleAspSerLysGlyAenLeuAenSerAenAlaGlnIleGln 288
QY      1162  -----TCATCTACTTCAACAGAA 1179
Db      289  SerAlaSerThrIleAenLeuThrAenGlyThrLeuAenAenThrThrGlyThrVal 308
QY      1180  GATCCAGTAGTAGTATTTTCC-----AGAAATACTCGGTAGAGTTT---GAT 1224
Db      309  ThrSerValGlyThrIleSerLeuAenThrAenLysAenThrIleValAenThrArgAla 328
QY      1225  GGGAACTGACCCGAGTAGGA----- 1245
Db      329  GlyAenIleSerThrMetGlyAspIleTyrValAenSerGlyThrIleAspAenThrAen 348
QY      1246  GGAGGATTTACTCTACCGGAACGTTGCTTCTCGAATAATGAAACCTTGTCTTC 1305
Db      349  GlyLysLeuAlaAlaAlaGlyMetLeuAlaValAspThrAenAenAlaThrLeuIleAen 368
QY      1306  AACAAATGTTCTCTCTCTGTT-----TACATTGCTGCTAGCAACCAACA 1350
Db      369  SerGlyLysGlySerSerValGlyIleGluAlaGlyLeuValAlaLeuLysThrGlyThr 388
QY      1351  AGTGGACAGCGCTTCTAATACGAGTAATAATACGAGATCGAGGAGCTATCTTCTGTAA 1410
Db      389  -----LeuAenAenSerAenGlyGlnIleArgGlyGlyTyrValGlyLeuGlu 404
```

1411 AATGGTGGCAGGAGGATCCAAATAACTCTGGATCAGTTTCTCTTGATGGAGGGAGTA 1470
Db SerAlaAlaLeu-----AsnAsnAsnAsnGlyAspIleGlnThrThrGlyAspIleAla 422
1471 GTTTCTTTAGTACGAATGTAGCTCTGGGAAAGG----- 1506
Db IleIleSerAsnGlyAsnValAspAsnAsnLysGlyLeuIleArgSerSerThrGlyHis 442
1507 -----GGAGCTATTATTATGCAAAAAGCTCTCGTTGCTAACTGTGGC 1548
Db IleValIleGlyAlaAlaGlySerValAsnAsnGlySerThrLysThrAlaAspThrGly 462
1549 CCTGTACAAATTTTA-----AGGAATATC 1572
Db SerSerAspSerLeuGlyIleAlaAlaAspThrGlyValGluIleGlyAlaAsnAsnIle 482
1573 GCTAATGATGGTGGAGCGATTTATTAGGAGAACTCGGAGAGCTCAGTTTATCTGCT--- 1629
Db AsnAsnAsnGlyGlnIle-----AlaSerAsnGlyAsnValSerLeuSerSerTyr 500
1630 -----GATTATGGAGATATTATTTTCGATGGGAAT-----CTTAAAGA 1668
Db SerThrIleAspAspTyrAlaGlyLysIleLeuSerAsnSerLysValIleIleLysGly 520
1669 ACAGCCAAAGAGATGCTGCCGATGTTAATGGCGTAACGTGTCTCACAAGCCATTTGG 1728
Db SerSerLeuArgAsn-----AspThrGlyGlyIleSer---GlyLysGlnIleGly 537
1729 ATGGGATCGGAGGGAATAACGACATTAAGAGCTAAAGCAGGCGATCAGATCTCTTT 1788
Db ValAlaValGlyGlySerLeuThr----- 545
1789 AATGATCCCATCGAGATGCAACCGAAATAACCGAGCGAGCTCTTCCAAACTTCTA 1848
Db -----AsnAsnIleGlyValIleSerSerGlu----- 554
1849 AAAATTAACGATGGTGAAGGATACACAGGGATATT---GTTTTGCTAAT----- 1896
Db AsnHisGlyGlyPheMetMetGlyGlnAsnIleThrMetGluSerMetSerGlyValAsn 566
1897 -----GGAGCAGTACTTTGTACCAAAATGTTAGATAGACAA----- 1935
567 AsnHisGlyGlyPheMetMetGlyGlnAsnIleThrMetGluSerMetSerGlyValAsn 586
1936 -----GGAAGGATGTTCTTGTGAAAGGCAAAATATCAGTG---AATTCTCTA 1983
Db AsnAsnThrAlaLeuIleValAlaSerLysLysLeuLysIleAsnAlaArgGlySerIle 606
1984 AGTCAGACAGGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCA 2043
Db GluAsnArgAspGlyAsnAsnPhe-----GlyAsnAlaTyrGlyLeuTyrPheGly 623
2044 CAACACCAACACAGCTCTCGCGTAACTACATCAGTTGATCAGCTTTCCAACTGTGATTG 2103
Db MetProGlnGln-----ThrGlyGlyMetValGlyLysGluGlyIleGluLeu 639
2104 TCTCTTTCTCTTTGTTAGCAACATGCA-----GTTACGAATCTCTCTACCAAT 2154
Db SerGlyGlnAsnIleTyrAsnAsnAsnSerArgLeuIleAlaGluAspGlyProLeuThr 659
2155 CTTCCAGCGCAA-----GATTCTCATCTCGAGTCATTTGGTAGCAACAATGCTGCTG 2205
Db LeuGlnAlaGlnAsnThrPheAspAsnThrArgAlaLeuValThrSerGlyAlaAspAla 679
2206 TCTGTTACAAATTAGTGGGCTATCTTTTGGAGATTGATGATACAGCTTATGATAG 2265
Db SerIleGlnValGlyGly-----ThrTyrTyrAsnAsn 690
2266 TAT-----GATTGGCTAGGTTCTAATCAAAAATCAATGTCCTGAAATTACAG----- 2313
Db TyrAlaThrThrTyrSerAlaGlyAsnLeuAspIleAspAlaThrThrLeuGlnAsnSer 710

2314 ---TTAGGACTAAGCCCCCAAGCTAATGCC-----CCATCAGATTTCGACTCTA 2358
Db SerSerGlyThrMetIleAspAsnAsnAlaThrGlyPheIleAlaSerAspLysAsnLeu 730
2359 GGG-----AATGAGATGCTAAGTATATGGCTAT-----CAAGGAAGCTGG 2397
Db SerLeuGluValValAsnSerLeuThrAsnTyrGlyTyrIleSerGlyLysGlyAspVal 750
2398 AAGCTGTGGTGGGATCCTAATACAGCAAAATATGCTCCT---TATACTCTCAAGACTACA 2454
Db AspVal-----ThrValAsnAsnGlyAsnLeuTyrAsnArgAsnThrIle 765
2455 TGGACTAAACTGGGTATATCTCTGGGCTGAGCGAGTAGCTCTTGGTTCCAAATAGT 2514
Db AlaAlaGluLysGlyLeuAsp-----IleAlaAlaLeuAsnGlyIleGlu 780
2515 TTATGGGATCCATTTTA-----GATATA---CGATCTGGCATTCAGCAATTCAA 2562
Db AsnTyrLysAspIleSerAlaGlyGlyAspLeuThrMetAsnThrAsnArgHisValThr 800
2563 CCAAGTGTGGATGGGCGCTCTTATTCGAGGATTTATGGTTCCTGGAGTTTCGAATTC 2622
Db AsnAsnSerAsnSerAsnMetValGlyGlnAsnIleValIleAsnAlaValAsnAspIle 820
2623 -----TTCTATCATGACCGCAT-----GCTTTAGGTCAGGA 2655
Db AsnAsnArgGlyAsnIleValSerAspAlaAspLeuAsnValThrThrLysGlyAsnLeu 840
2656 TATCGGTATAT-----AGTGGGGTATTCTCTTAGGAGCAAACTCTCTACTTT 2703
Db TyrAsnTyrLeuTyrMetValGlyTyrGlyAspIleAlaLeuSerAlaAsnSerValAla 860
2704 GATCATCATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGTCTTAAAGATTATGTA 2763
Db AsnAsnAsnAlaThrIleGluAlaThrGlyAspLeuIleIleAspSerLysGlyAsnVal 880
2764 ---GTGTGCTCTCCAATCATCATCTGTGCATAGATCGTT----- 2802
Db GlyAsnAsnArgGlyAsnLeuHisAlaLeuAsnGlyValLeuSerValLysGlyAsnAsn 900
2803 -----TATCTATCTACCCCAACAAAGCTTTTATGTGA 2832
Db LeuAsnAsnAspAsnGlyGluIleArgGlyTyrGlyAspValThrLeuAlaLeuThrGly 920
2833 TCCTATTGTTGGAGATGCGTTTATCCGTGCTAGCTACGGGTTGGGAATCAGCATATG 2892
Db AsnTyr-----AspSerTyr----- 925
2893 AAAACCTCATATACATTTGACAGAGAGAGCGATGTTCTGTTGGATAATAAC----- 2943
Db LysGlySerLeuThr---SerGluThrGlyAspValThrLeuThrAlaAsnIleValAsp 944
2944 -----TGTCTGGCTGGAG-----ATTGA 2964
Db AsnAlaTyrGlyLeuIleAlaGlyGluAsnValSerValAspAlaLysSerThrIleTyr 964
2965 GCGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTGTAATGAGTGTGCTCTTC 3024
Db AsnAsnThrAlaLeuIleAlaAlaAsnLysLysLeuValIleAsn----- 979
3025 GTGCAAGCTGAGTTTCTTATGCCCATCATGAACTCTTTTACAGAGGAGCGCATCAAGCT 3084
Db AlaGlyGlyAsnLeuGluAsnArgAspGlyAsnAsnPheLeuArg----- 994
3085 CGGGCATTCAGAGCGGACATCTCCATAAATCTATCAGTTCTCTGTTGGAGTGAAGTTGAT 3144
Db -----AsnAsnGlyAlaLeuPheGlyIleThrAspAsnValGly----- 1007
3145 CGATGTTCTAGTACATCTCTAATAATATATAGCTTTTATGGCGGCTTATATCTGTATGCT 3204
Db ----- 1007
3205 TATCGCACCATCTCTGTGTTACTGAGACAAGCGCTCTATCCATCAAGAGACATGGACAACA 3264

```

; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1245
; LENGTH: 3132
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-1245

Alignment Scores:
Pred. No.: 1,22e-14 Length: 3132
Score: 270.50 Matches: 300
Percent Similarity: 34.1% Conservative: 192
Best Local Similarity: 20.8% Mismatches: 560
Query Match: 3.4% Indels: 390
DB: 11 Gaps: 67

US-10-701-844-1 (1-4435) x US-11-087-099-1245 (1-3132)
QY 505 TTAACCTGATCAATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTTCT 564
Db 505 TTAACCTGATCAATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTTCT 564
QY 35 IleMetLeuSerPheProGlyGlnAlaLeuAlaGlyProSerGlyGlnValIle 53
Db 35 IleMetLeuSerPheProGlyGlnAlaLeuAlaGlyProSerGlyGlnValIle 53
QY 565 GCAGGAGAGTTAAACATTA---AAAAATCTTGACAATTTCTATTGCAGCTTTGCCTTTAAGT 621
Db 565 GCAGGAGAGTTAAACATTA---AAAAATCTTGACAATTTCTATTGCAGCTTTGCCTTTAAGT 621
QY 54 AlaGlyGlnValThrIleAspArgLeuAlaLeuAspThrIle 67
Db 54 AlaGlyGlnValThrIleAspArgLeuAlaLeuAspThrIle 67
QY 622 TGTTTTGGGAACCTTATTAGGAGTTTACTGTTTATGGGAGAGGACACTCGTTTGACTTTC 691
Db 622 TGTTTTGGGAACCTTATTAGGAGTTTACTGTTTATGGGAGAGGACACTCGTTTGACTTTC 691
QY 67 67
Db 67 67
QY 682 GAGACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741
Db 682 GAGACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741
QY 68 ---IleAsnGlnAlaThrAspLysAlaIleAsnTrpLysAsnPheAsp 83
Db 68 ---IleAsnGlnAlaThrAspLysAlaIleAsnTrpLysAsnPheAsp 83
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATATCTCTTTCCAAATTGGCAATTTCATTCTTCCGTA 801
Db 742 TTTTACTATTGAGGGTTTTAAAGAAATATCTCTTTCCAAATTGGCAATTTCATTCTTCCGTA 801
QY 84 84
Db 84 84
QY 802 CTGCGCT---GCTCCAAACGACTAATAAGGTAGCCAGACTCCGACGACAAACA 849
Db 802 CTGCGCT---GCTCCAAACGACTAATAAGGTAGCCAGACTCCGACGACAAACA 849
QY 93 IleGlnProAspValGlySerIleAlaLeuAsnArg 107
Db 93 IleGlnProAspValGlySerIleAlaLeuAsnArg 107
QY 850 TCTACACCGCTCAAT---GGTACTATTATTCTTAAACACAGACTCTTTTGTACTCAAT 903
Db 850 TCTACACCGCTCAAT---GGTACTATTATTCTTAAACACAGACTCTTTTGTACTCAAT 903
QY 108 AsnLeuProThrAsnIleLeuGlyGlnLeuAlaAsnGlyArgLeuPheIleLeuAsn 127
Db 108 AsnLeuProThrAsnIleLeuGlyGlnLeuAlaAsnGlyArgLeuPheIleLeuAsn 127
QY 904 AATGAGAAGTCTCTCAATCTATATGTAATTTAGTCTCTGGAGATGGGGGAGCTATAGTCT 963
Db 904 AATGAGAAGTCTCTCAATCTATATGTAATTTAGTCTCTGGAGATGGGGGAGCTATAGTCT 963
QY 128 128
Db 128 128
QY 964 AAGAGCTTAACCGTTCAAGAAATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAA 1023
Db 964 AAGAGCTTAACCGTTCAAGAAATTAGCAAGCTTTGTCTTCCAGAAATACTGCTCAA 1023
QY 142 AlaGlyLeuLeuAlaThrLeuSer 157
Db 142 AlaGlyLeuLeuAlaThrLeuSer 157
QY 1024 GCTGATGGGGAGCTTGTCTCAAGTATGACCTACCACTTTC---TCTCTATGGCTACGAG 1077
Db 1024 GCTGATGGGGAGCTTGTCTCAAGTATGACCTACCACTTTC---TCTCTATGGCTACGAG 1077
QY 158 SerGlyLysTyrAlaPheGluLysValProGlyLeuLysAspAlaValGluAsnLys 177
Db 158 SerGlyLysTyrAlaPheGluLysValProGlyLeuLysAspAlaValGluAsnLys 177
QY 1078 GCTCCTATT---GCCCTTTAGCGAATGTTGCGAGGAGTAAGA 1116
Db 1078 GCTCCTATT---GCCCTTTAGCGAATGTTGCGAGGAGTAAGA 1116
QY 178 GlyGluIleHisIleGlyGluGlyPheAlaPheLeuVal---AlaProGlyValArg 196
Db 178 GlyGluIleHisIleGlyGluGlyPheAlaPheLeuVal---AlaProGlyValArg 196
QY 1117 GGGGGAGGAGTCTGCTGTTTCAGGATGGGACAGGAGTGTCTATCTACTTCAACA 1176
Db 1117 GGGGGAGGAGTCTGCTGTTTCAGGATGGGACAGGAGTGTCTATCTACTTCAACA 1176
QY 197 AsnGluGlyLeuIleAlaLysLeuGlySerValAlaLeuAlaSerGlyAsp 214
Db 197 AsnGluGlyLeuIleAlaLysLeuGlySerValAlaLeuAlaSerGlyAsp 214
QY 1177 GAAGATCCAGTAGTAAGTTTTCAGAAATATCTCGGAGTAGTTGATGGGACGTAGCC 1236
Db 1177 GAAGATCCAGTAGTAAGTTTTCAGAAATATCTCGGAGTAGTTGATGGGACGTAGCC 1236
QY 215 ---LysMetThrIleAspPheHisGlyAsp 223
Db 215 ---LysMetThrIleAspPheHisGlyAsp 223

1008 -----GlyIleValGlyLysGluGlyValThrLeuSerAlaGlnAsnValThrAsnAsn 1025
QY 3265 GATGCTTTTCAATTAGCAAGACATGGA-----GTTGTGGTTAGAGGATCTATG--- 3312
Db 1026 AsnSerIleIleAlaGluAlaGlyProLeuAsnLeuLeuSerArgGlyThrLeuAsp 1045
QY 3312 ----- 3312
Db 1046 AsnThrArgAlaLeuLeuSerSerGlyAlaAspAlaIleIleArgAlaAlaGlyThrPhe 1065
QY 3313 -----TATGCTTCTCTAACAAGT-----AATATAGAAGTATAT----- 3345
Db 1066 TyrAsnAsnTyrAlaThrThrThrSerAlaGlyAsnLeuAspValThrAlaAlaSerLeu 1085
QY 3346 -----GGCCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTGAGT 3396
Db 1086 AsnAsnAlaSerAspGlyArgLeuGluAspAsnThrAlaThr-----GlyValIle 1102
QY 3397 GCAGGAGTAGAGTCCGGTCTTCTAAATAATATGTTTATAGATGTTAAGTGTAGCGTCT 3456
Db 1103 AlaSerAspLysAsn-----LeuAsp----- 1113
QY 3457 TTTTCTTTGAGATCTACATCATTTTCTTTTGTGTTGTTGTTCTTCTTATCTGATGGA 3516
Db 1114 AsnSerValThr-----AsnTyrGly 1120
QY 3517 TTCGCGAGCTCTCTCAAGTGTAAACGCTTAATGTAAACCATCTCTTTTAAAGGAGACGAT 3576
Db 1121 TrpIleSerGly-----LysGlyAsp--- 1127
QY 3577 GTTTACTGTAATGGAGACTGCGGCTTTTGTCTCAATGCTATGTCAGGAGCTGAAGAAGTTCG 3636
Db 1128 -----ValHisPheAsnValLeuLysGlyThr 1136
QY 3637 ATTATCTCAGCTAAT---GGCGACAAT---TTAACGATATTACCGACACAAACCAT 3684
Db 1137 LeuTyrAsnArgAsnAlaIleAlaIleAspAsnAlaLeuThrIleAsnAlaLeuAsnGly 1156
QY 3685 ACATTATCATTTACAGATTTCTCAAGGCCAGTCTTCCAAATTTATGCTTTCATTTCAGCA 3744
Db 1157 ValGluAsnPheLysAsp-----IleValAla 1165
QY 3745 GAGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTGC 3804
Db 1166 GlyThrAlaLeuThrIleAspThr-----GlnLysTyrValThrAsn 1179
QY 3805 GGAGAAAAGGAATGATCTCCGGGAAAACCGTGAGTATTTCCGAGCGACGCGAAGTGATT 3864
Db 1180 AsnSerAsnSerAsnMetLeuGlyGlnThrIleAlaIleAsnAlaValAsnAspIleAsn 1199
QY 3865 TTCTGGGATAACTCCGTTGGGGTATTTCTCTTTATCTACTGTGCGCAACCTCATCAACT 3924
Db 1200 AsnArgGlyAsnIleValGlyAspTyrSerLeuGlyValLysThr----- 1215
QY 3925 CCGCCTGCTCCACAGTAGTATGATGTCGGAAGGGTCTATTTTCTGTAGAGACTAGT 3984
Db 1216 -----GlyAsnIleTyrAsnTyrLeuAsnMet 1224
QY 3985 TTGCGAG-----ATCTCAGCGCTCAAAAAGGGTCTATGTTCCGATATAAT----- 4029
Db 1225 LeuSerTyrGlyValAlaGlyValSerAlaAsnLysValThrAsnSerGlyLysAspAla 1244
QY 4030 -----GCCCGGAATTTCCGACACAGATTTTTCGAGGTAAGAATAATAATGCTGGT 4080
Db 1245 ValLeuGlyGlyPheTyrGlyLeuAlaLeuGluAlaAsnGluThrAspAsnThrGly 1263

RESULT 21
US-11-087-099-1245
; Sequence 1245, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al..
```

```

QY 1237 CGAGTAGGAGGAGGATTTACTCTCCAGCGGAACGTTGCTTCTTGATTAATGTAARACC 1296
Db      |||::: |||||
224      GlyLeuValSerTyrGlu----- 229
QY 1297 TTGTTTCTCAACAATGTTGCTTCTCTCTGTTTACATTGCTTAAGCAACCAACAGTGG 1356
Db      |||::: |||||
230      ---ValThrGlyAlaValAlaGluLysValLeuGlyAlaAspGlyGlnProLeuGluAla 248
QY 1357 CAGGCTTCTAATACAGTAATAATTAACGAGATGAGGAGCTATCTTCTGTAAGATGTT 1416
Db      |||::: |||||
249      LeuValSerAsnThrGlyThrIleGlnAlaAspGlyGlyAsnValValMetGln---Gly 267
QY 1417 GCGCAAGCAGGATCC-----AATAACTCTGGATCAGTTTCTCTTT--- 1455
Db      |||::: |||||
268      HisSerAlaGlyAspValPheSerSerValValAsnAsnGluGlyIleIleGlnAlaGln 287
QY 1456 -----GATGAGAGGGAGTAGTTTTC 1476
Db      |||::: |||||
288      SerLeuGlnLysLysAsnGlyLysIleTyrLeuGlyGlyAspAspGluGlyGlyIleVal 307
QY 1477 TTTAGTAGC-----AATGTAGCTCTGGAAA-----GGGGAGCTATT 1515
Db      |||::: |||||
308      SerAsnSerGlyIleLeuAspValSerAlaAlaGluProAsnSerAlaGlyGlyGluVal 327
QY 1516 TATGCCAAAAGCTCTCGTTGCTTAAGTGTGGCCCTGTACAATTTTAAAGNAATCGCT 1575
Db      |||::: |||||
328      ValMetThrGlyHisPheValGlyAsnGluGlyValIleGluAlaLysGlySerGlySer 347
QY 1576 AATGATGTGGAGCAATATTATTAGGAGAACTCGA----- 1611
Db      |||::: |||||
348      GlyAspGlyGlyGlnValThrLeuIleSerAlaGlnHisThrLeuThrThrValAsnSer 367
QY 1612 -----GAGCTCAGTTTATCTGCTGATTTAGGAGAT----- 1641
Db      |||::: |||||
368      ValIleAspValSerGlyGlyAlaAspSerGlySerGlyGlyValValGlnIleArgSer 387
QY 1642 -----ATTATTTTCATGGGAATCTTAAAGAACAGCAAGAGAGATCTCCCGAT 1692
Db      |||::: |||||
388      AlaAsnA-gAlaThrPheAspGlyValValLysAlaAspAlaGlyGlnAsnGlyGlyAsp 407
QY 1693 GTTAATGGGTCACTGTCTCACAAGCCATTTGCGATGGGATCGGAGGAGGAAAATAACG 1752
Db      |||::: |||||
408      GlyGlyLeuValAspValSerSerGluGly---GluValGlnLeuIleGlyLysValVal 426
QY 1753 ACATTAAAGCTAAAGCAGGCGATCAGATCTCTTTAATGATCCC-----ATCAGATG 1806
Db      |||::: |||||
427      GlyLeuAlaProArgGlyArgAsnAlaLysLeuValIleAspProLysHisIleGluIle 446
QY 1807 GCARAACGGAAATAACAGCCAGCGAGCTTCCAACTTCTAAAATTAACGATCGTGA 1866
Db      |||::: |||||
447      ---AsnAspAlaGlyGlyAlaTyrAsnSerThrThrValGlnPheAspAspSerGln 465
QY 1867 GGATACACAGGGGATATGTTTGTCTAATGGAAGCAGTCTTGTACCAAAATGTTACG 1926
Db      |||::: |||||
466      SerGlyThrThrValIleThrAlaAla-----SerIleAsn 477
QY 1927 ATAGAGCAAGGAGGATTTCTTCTCGTGAAGGCAAAATTTATCAGTGAAT----- 1977
Db      |||::: |||||
478      AlaGlnValAlaAspValValLeuGlnAlaAsnThrAspIleThrValThrThrAlaIle 497
QY 1978 TCTCTAAGTCACAGAGTGGGAGTCTGTATATGGAAGCTGGGAGTACA----- 2025
Db      |||::: |||||
498      AsnIleAlaAsnAlaGlyThrGlyLeuThrLeuGlnAlaGlyArgSerIleLeuIleAsn 517
QY 2026 TGGGATTTTGTAACTCCACAACCAACCAACAGCCCTCGCGCTTAATCAGTTTGAATCAG 2085
Db      |||::: |||||
518      AlaAspLeuThrThr-----AsnAsnGlyValValThr 528
QY 2086 CTT---TCCAACTCGATTTGCTCTTTCTCT----- 2115
Db      |||::: |||||
529      LeuThrAlaAsnAspHisThrAlaLeuSerAlaAspArgAspAlaGlyValAlaSerIle 548
QY 2116 TTGTTAGCAACAATGCAGTTACGAAT-----CCTCTTACCAATCTCCAGGCA 2166

```

```

Db      |||::: |||||
549      AlaMetAlaSerGlyThrThrIleAsnAlaGlySerAlaAlaValAsnMetThrIleAsp 568
QY 2167 GATTCTCATCTGAGTCAGTCATTTGGTAGCACAACTCTCTGGTTCTGTGTACAAATTAGTGGGCT 2226
Db      |||::: |||||
569      AlaSerGlyHisGlyLeuMetGlyAspIleThrAlaTyrAsnValThrSerSerGlyAla 588
QY 2227 ATCTTTTGGAGGATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTTAAT 2286
Db      |||::: |||||
589      ValAlaLeuAspSerAlaGlySerIleValAsnGlyProAsnPheAlaSerGlySer 608
QY 2287 CAAAAAATCAATCTGCTGAAATTCAGTTAGGAGCTAAGCCCCAGCTAATATGCCCATCA 2346
Db      |||::: |||||
609      ThrSerIleGlyLeu-----ThrAlaAspAsn 617
QY 2347 GATTGTACTTAGGAATGAGATGCTTAAGTATGCTATCAAGAGCTGAAGCTGG 2406
Db      |||::: |||||
618      AspValThrValGlySerAlaLeuThr-----ThrGlySerLeuThrIleAla 633
QY 2407 TGGGATCCTAATACAGCAATAATGCTTATACT-----CTGAAA 2448
Db      |||::: |||||
634      AlaGlyArgSerValAlaLeuAsnSerAsnValThrThrThrSerThrValAsnIleLys 653
QY 2449 GCT-----ACATGACTAAACTGGGTATAATCTCTGGCCTGAGCGAGTA 2493
Db      |||::: |||||
654      AlaAsnAspSerThrGlnThrLeuAlaAsnArgGlySerGlyAspGlyAlaIleSerMet 673
QY 2494 GCTTCTTTGGTCCAAATAGTTTATGGGATCCATTTTATAGATATACGATTCGCGCATCA 2553
Db      |||::: |||||
674      AlaAla-----GlyThrThrLeuAsnGlyGlySerAla----- 684
QY 2554 GCAATTCAAGCAAGTGTGGAT-----GGGCGCTTATTGTCGAGGATTTGGTTTCT 2607
Db      |||::: |||||
685      AlaIleThrLeuThrValAspSerThrAsnAlaSerThrAlaGlyGlyIleThrIleAsp 704
QY 2608 GGAGTTTCG-----AATTCTTCTATCATGCGCATGCTTTTAGTTCAGCGAGATAT 2658
Db      |||::: |||||
705      LeuValThrThrAspGlyAsnLeuThrValSerSerHisAspAlaIleThrGlu----- 722
QY 2659 CGGTATATTAGTGGGGTTATTCCTTAGAGCAAACTCTCTAC-----TTTGGATCATCG 2712
Db      |||::: |||||
723      -----IleAspAlaAspSerGluValAspLeuArgGlyAsp 734
QY 2713 ATGTTTGTCTAGCATTTTACCGAAGTATTTGCT-----AGA 2748
Db      |||::: |||||
735      IleLeuSerLeuThrThrThrAspThrAsnGlyAlaTyrIleGlyAlaProGlyThrSer 754
QY 2749 TCTAAAGATTTAGTGTGTGCTTCCCAATCATCTGCTTGCA-----GGA 2796
Db      |||::: |||||
755      ThrThrAspTyrLeuGluValHisAlaLysThrArgValAspValThrThrAsnAsnGly 774
QY 2797 TCCGTTTATCTATCTPACCAACAAGCTTTATGTGATCCTAT---TTGTTCCGA----- 2847
Db      |||::: |||||
775      SerValAlaLeuAlaSerProGlyTyrLeuSerGlyThrValAspTyrPheGlyAlaAla 794
QY 2848 -----GATGCTTTATCCGTGCTAGCTACGGGTTT 2877
Db      |||::: |||||
795      ProThrAspAspIleGlyThrThrAspIleAspGluSerLysIleValAsnLeuAsnLeu 814
QY 2878 GGGATCAGCATATCAAAACCTCATATACATTTGCGAGGAGGAGCGATGTTCTGTTGGGAT 2937
Db      |||::: |||||
815      GlyThrValAsnValGlyThrAla---ThrPheMetLeu-----ValValThrAsp 830
QY 2938 AATACTGTCTGGCTGGAGAGATTGGAGCGGATTTACCGGATTTGATTTACTCTCACTAAG 2997
Db      |||::: |||||
831      GlyGlyIleLeuAspGlyLeuAlaGlyThrGlyThrAsnIleThrAlaGlySerAlaAsn 850
QY 2998 CTCATTATTGAATGAGTTGCGTCCCTTCGTCGAGCTGAGTTTCTTATGCC---GATCAT 3054
Db      |||::: |||||
851      LeuThrThrAsnGlyGlyGluSerValValAlaGlyAlaIleValLeuGlyAsnAspGly 870
QY 3055 GAATCTTTTACAGAGGAAGGAGATCAAGCTCGGCAATTCAGAGC-----GGACATCTTC 3108
Db      |||::: |||||

```

```
Db 871 ThrArgGlyGlySerIleGlyAspAlaThrHisAlaLeuLeuThrGluLeuGlyIleLeu 890
Qy 3109 CTAATCTATCAAGTTCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCTAAT 3168
Db 891 ThrAlaSerSerAsnGlySerValTyrIleAspGluAlaSerGlyIleLeuIleAsn 910
Qy 3169 AAATATAGCTTT-----ATGGCGGCTTATATCTGTGATGCTTATCGCACCCATC 3216
Db 911 SerIleAlaAsnGlnIysGlyValThrAlaIysValAsnSerSerTyrAlaValVal 930
Qy 3217 TCTGCTACTGAGACACGCTCTCTATCCATCAAGAGACATGACACAGATCGCTTTCAT 3276
Db 931 -----ValThrThrLeuAspGlyGlnThrGluThrGlyThrAsnAsnValThrVal 947
Qy 3277 TTAGCAAGACATGGAGTGTGGTTAGAGATCTGTATGATGCTTCTCTAACAAGTAAATA 3336
Db 948 ThrAlaAla---GlyAspIleIleIleGlyThrAlaSerAlaThrSerThrValThrLeu 966
Qy 3337 GAAGTATATGGCCATCGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGA 3396
Db 967 ThrSerThrAlaGlyAsnIleTyrAspAsnAsnAspAlaThrAsn----- 981
Qy 3397 GCAGGAAGTAGAGTCGGTTCATAAATAATGCTTAGATAGTAAAGTGTAGCGATGCCT 3456
Db 981 ----- 981
Qy 3457 TTTTCTTGAGATCTACATCATTTGTTTGTAGTGTGTTGTTGTTCTTATTCGTATGGA 3516
Db 981 ----- 981
Qy 3517 TTTCGGAGCTCTCTCAAGTGTAAAGCCTAATGTAACCACTCTCTTTTAAGGAGAGCAT 3576
Db 982 -----AsnIleLeuAlaIysAsnValVal----- 989
Qy 3577 GTTACTTCTGAAGTGGAGTGGCTTTTGTCAATGCTCTATGTCAGAGCTGAAGAAGTTCG 3636
Db 990 -----LeuAsn-----AlaGlyLeuAsnMetGly--- 997
Qy 3637 ATTATCTAGTAATGCGACAAATTAACGATTACCGGACAAACCATCATCATTCATT 3696
Db 998 -----LeuSerGlyAspValLeuLeuLeuThrGluSerAlaThrLeuThrAla 1014
Qy 3697 ACAGATTCTCAAGGCCAGTCTTCAAAATATTGCTTCAATGCTTCTTCAAGCAGAGACAT 3756
Db 1015 ThrAsn-----GlyGlyValTyrPheAlaMetGlyValProSerThrValAlaSerVal 1032
Qy 3757 ACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAATGCTTTCTTGGCGAGAAAAGGA 3816
Db 1033 ThrAlaGlyGlyHisAspValValIleAsnAlaSerAspGlyValLeuSerLeuGly 1052
Qy 3817 ATGATCTCC-----GGGAAACCGTGAGTATTTCCGGAGCAGCGGAGTGATTTCTGG 3870
Db 1053 MetValSerAlaValGlySerValThrIleAspGlyThrGlySerIleLeuAsp 1072
Qy 3871 GATACTCCGTGGGTATCTCTTATCTACTGTGCCACCTCATCATCACTCCGCT 3930
Db 1073 AsnAsnGlyAlaThrAlaAlaIleAlaAsnSerAlaSerIleIysSerAspIysGly 1092
Qy 3931 GCTCCAACAGTATGATGTCGGAAGGCTCTATTTTCTGTAGAGACTAGTGTGGAG 3990
Db 1093 AlaGlyThrAlaSerAspAlaLeuGluThrThrValAspThrLeu-----AlaValGlu 1110
Qy 3991 ATCTCAGGCGTCAAAAGGGGTCATGTTTCGATAAT-----AAT 4029
Db 1111 IleThrGlySerGlyIysSerPheTyrIleAspGluSerAspAlaLeuThrSerIleAsn 1130
Qy 4030 CGCGGGAATTCGGAACAGTTTTTCGAGGTAAGATAATATATATGCTGCTGCTGGAGGC 4089
Db 1131 Ala-----LysValAsnAsnGlySerThrAsnLeuAsnPheThrGlyGlySer 1146
Qy 4090 AGTGGGTTCGCTACACCATCAAGTACGACTTTTACAGTTTAAACACTGTAAGCGAAAGT 4149
Db 1147 PheAlaPheAsnAlaThrThr-----GlyAlaPheSer-SerThrGlyValGlySerVa 1164
```

```
Qy 4150 TTCTTTACAGATAACGTAGCTCTTTCGGAGGGGAGTGGTTTATAAAGGCATTGTGCT 4209
Db 1164 lThrPheGluAsnThr-----GlyGlyGlyValAlaIleGlyThrValThrAl 1180
Qy 4210 TTCAAAGACAAATGAAGGAGCATATTTCTCCGAGGGAACACAGCATAGCATTTAAG 4269
Db 1180 a-----ThrGlyGlySerAlaThrIleThrAlaSerThrAlaIleThrAsp----- 1195
Qy 4270 GATTTCTGCTCTACTAATCAGGAT-----CAGAAATACGAGACAGGAGG 4314
Db 1196 -AlaThrSerAlaIleThrAlaAspThrValValLeuThrAlaGlyThrSerIleGlyAl 1215
Qy 4315 CGGTGGAGGAGTATTTGCTCTCCAGATCATTTCTGAAAGTTTGAAGGCAATAAAGTTTC 4374
Db 1215 aserGlySerThrIleIysSerThrAlaThrThrLeuThrTyrLeuAlaSerAlaGlySe 1235
Qy 4375 TATT 4378
Db 1235 rfile 1236

RESULT 22
US-11-052-554A-91
; Sequence 91, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 91
; TYPE: PRT
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-91

Alignment Scores:
Pred. No.: 1,61e-13 Length: 2902
Score: 257.50 Matches: 302
Percent Similarity: 31.4% Conservative: 171
Best Local Similarity: 20.1% Mismatches: 472
Query Match: 3.3% Indels: 560
DB: 11 Gaps: 79

US-10-701-844-1 (1-4435) x US-11-052-554A-91 (1-2902)
Qy 403 TTCTTTTTCATGATCTTAGCTTATCTTGTGCTCTTTTAAATGGGGGATATGCA 462
Db 119 TyrPheLeuSerGlyLeuTyrAsnTyrThrGlyGlyLeuTyrAsnGlyGlyAsnLeuAsp 138
Qy 463 GCAGAAATC-----ATGGTTCCTCAAGGAATTTACGATCGGAGACGTTAACT 510
Db 139 IleGluLeuGlySerAsnAlaThrPheAsnLeuGlyAlaSerSerGlyAsnSerPheThr 158
Qy 511 GTATCATTTCCCTATCTGTTATAGGATCGGAGTGGGACTACT-----GTTTTTCT 564
Db 159 SerTrpTyr-----ProAsnGlyHisThrAspValThrPheSer 171
Qy 565 GCAGAGAGTTAACTATTAATAATCTTGACAAATCTTATTCAGCTTTTGCCTTTTAAGTTGT 624
Db 172 AlaGlyThrIle-----AsnValAsnAsnSerValGluVal----- 183
Qy 625 TTTGGGAACTTATTAGGAGTTTACTGTGTTTATAGGAGAGGA---CAC----- 669
Db 184 ---GlyAsnArgValGlySer-----GlyAlaGlyThrHisThrGlyThrAla 198
```



```

, , TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
, , FILE REFERENCES: 30853/40359A
, , CURRENT APPLICATION NUMBER: US/11/052,554A
, , CURRENT FILING DATE: 2005-02-07
, , PRIOR APPLICATION NUMBER: US 60/589,227
, , PRIOR FILING DATE: 2004-07-20
, , PRIOR APPLICATION NUMBER: IN 173/DEL/2004
, , PRIOR FILING DATE: 2004-02-06
, , NUMBER OF SEQ ID NOS: 763
, , SOFTWARE: PatentIn version 3.3
, , SEQ ID NO 90
, , LENGTH: 3194
, , TYPE: PRT
, , ORGANISM: Helicobacter pylori J99
US-11-052-554A-90

```

Alignment Scores:		
Pred. No.:	1.7e-12	Length:
Score:	246.00	Matches:
Percent Similarity:	30.8%	Conservative:
Best Local Similarity:	19.3%	Mismatches:
Query Match:	3.1%	Indels:
DB:	11	Gaps:
		81
		3194

US-10-701-844-1 (1-4435) x US-11-052-554A-90 (1-3194)

Qy	421	CTAGCTTATTCCTGCTCTCTTTAAATGGGGGGGATATGACGAGAAATCATGTTCTCT	480
Db	231	IleThrTyraaGlyGlySerValaAenGlyGlyaAenPheGlyPheAaAenValaAenSer	250
Qy	481	CAAGGAATTTACGATGGGGAGACGTTAACTGTATCATTTCCCTATACTGTTATAGACGAT	540
Db	251	AenGlyala-----ThrThrIle-----	256
Qy	541	CCGAGTGGGACTACTGTTTTTTTCTGCAGAGAGTTAAACATTAATA-----	585
Db	257	--SerGlyValThrPheAaAenAenAenGlyalaLeuThrTyrrLysGlyGlyAenGlyIle	275
Qy	586	-----AATCTTGCACAAT-----	597
Db	276	GlyGlySerIleThrPheThrAenSerAenIleAenHisTyrrLysLeuAenLeuAenAla	295
Qy	598	TCTATTGCAGCTTTGCCCTTTAAAGTTGTTTGGG-----AACTTA	636
Db	296	AenSerValThrPheAaAenSerThrLeuGlySerMetProAenGlyAaenAlaAenThr	315
Qy	637	TTAGGAGTTTTACTGTTTTAGGGAGAGACACTCGTTGCACCTTCGAGAACATACGGACT	696
Db	316	IleGlyAenAlaTyrrIleLeu---AenAlaAenAenIleThrPheAaAenLeu-----	332
Qy	697	TCTCAAAATGGGGCAGCTCTAAGTAATACGCTGCTGATGGACTGTTTACTTATTCAGGGT	756
Db	333	ThrPheAenGlyGlyTrpPheValPheAaenArgSerAaspAlaHisValaAenPheGlnGly	352
Qy	757	TTTAAAGAAATTATCC-----TTTTCCAATTGCAATTCATTACTTGCCTGA	801
Db	353	ThrThrThrIleAaAenProThrSerProPheValAaenMetThrGlyLysValThrIle	372
Qy	802	CTGCCTGCTGCAACGACATAAAGGATAGCCAGACTCCGACG-----	843
Db	373	AenProAenAlaIlePheAenIleGlnAenTyrrThrProThrIleGlyAenAlaTyrrThr	392
Qy	843	-----	843
Db	393	LeuPheSerMetLysAaenGlyAenIleAlaTyrrAaspAaspValaAenAenLeuTrpAenIle	412
Qy	844	-----ACAACTCTACACCGTCTAAT-----	864
Db	413	IleArgLeuLysAenThrGlnAlaThrLysAaspAaenSerLysAaenAlaThrSerAaenAaen	432
Qy	865	-----GGTACTATTATTCT-----	879
Db	433	AenThrHisThrTyrrValThrTyrrAenLeuGlyGlyThrLeuTyrrHisPheArgGln	452

Db 1351 AsnGlyGlySerIleThrPheGlyLysAsnAsnLeuLeuTyrLeuHisGlyAsnPheAsn 1370
Qy 3601 TTTGTCAATGCTCTAT-----CGAGGAGCTGAAGAGGT 3633
Db 1371 AlaThrAsnIlePheLeuThrAsnAsnPheAsnValGlyAsnProAsnAlaGlyGlyGly 1390
Qy 3634 TCATTATCTCAGTCAATGCGGAC---AATTAAACGATTACCGGACAAACACATACATTA 3690
Db 1391 AlaThrIleAsnPheAsnAlaAspGluThrLeuSerAlaAspGlyLeuAsnTyrThrAsn 1410
Qy 3691 TCATTACAGATTCTCAAGGCCAGCTTCTCAA-----AATTAGCCTTCATT 3738
Db 1411 PheGlnThrValAlaMetGly-----LeuGlnThrSerAlaSerGlnHisSerTrpAla 1428
Qy 3739 TCAGCAGGAGACACTTACTCTG-----AGAGATTCTTCG--- 3774
Db 1429 AsnPheAsnSerLysLeuSerMetGluIleLysAsnSerAsnPheArgAspPheThrTrp 1448
Qy 3775 ---AGTCTGATGTTCTCGAAAATGTTCTTCGCGAGAAAGGAATGATCTCGGGAAA 3831
Db 1449 GlyGlyPheArgPheAsnSerGlyArgIleThrPheGluAsnThrPheSerGlyTrp 1468
Qy 3832 ACCGTGAGTATTCCGAGCAGCGGAA----- 3858
Db 1469 Thr---AsnIleAsnGlyAlaThrGluSerGlySerSerTyrValAsnMetValAlaAsn 1487
Qy 3859 -----GTGATTTCTGGGATACTCCGTGGG----- 3885
Db 1488 ThrAspLeuIlePheThrAspSerIleLeuGlyGlyGlyLeuArgTyrAspLeuLysAla 1507
Qy 3885 ----- 3885
Db 1508 AsnAsnIleIlePheAsnAsnThrGlnMetValValAspValSerLysAsnValAsnGln 1527
Qy 3886 -----TATTCTCTTATCTACTGTGCGCAACCTCA 3915
Db 1528 SerSerLeuAsnGlyAsnValThrPheAsnHisSerArgLeuSerValLysProAsnAla 1547
Qy 3916 TCA-----TCNACTCCCGCTGCTCCACAGTGTAGTGTCTCGAAGGCTCT 3963
Db 1548 AlaIleAsnIleGlyGlyAspGlnThrGlnThrThrLeuGluAsnAlaSerSerLeuSer 1567
Qy 3964 ATTTTCTCTGAGAGACTAGTTTGGAGATCTCAGCGCTCAAA-----AAAGGGCTCATG 4017
Db 1568 PheTyrAsn---AspSerValAlaAsnPheAsnGlyThrAlaPheAsnGlyValSer 1586
Qy 4018 TTCGATAATATGCGCGGAATTCGGAACAGTTCGAGGTAGAGATAATAATAATGCT 4077
Db 1587 TyrLeuAsnLeuAsnProAsnAlaGlnValSerPheAsnGlnAlaAsnPheAsnAla 1606
Qy 4078 GGTGCTGAGCAGCTGGT-----TCCGCTACACATCA---AGTACGACTTT 4122
Db 1607 Asn-ValThrPheTyrGlyIleProLeuPheGlyLysThrProAsnPheGlyAsnSerVa 1626
Qy 4123 TACAGTTAAACAGTAAAGGCAAGTTCTTCTC-----ACAGATAACGTAGCCTC 4173
Db 1626 lArgLeuIleAsnPheLysGlyAspAlaLysPheAsnGlnAlaThrLeuAsnLeuArgAl 1646
Qy 4174 TTGCGGAGCGGAGTGGTTTATAAGGCATTTGTGCTTTTCAAAGCAATGAAGGA---GG 4230
Db 1646 aLysAsnIleHisLeuAsnPheGlnGlyAlaSerThrPheGluAsnAsnSerThrMetAs 1666
Qy 4231 CATATTCTTCGAGGGAACACAGCATACGATGATTAGGATTTCTTGCTGCTACTATCA 4290
Db 1666 nLeuAlaGluSerSerGlnAlaSerPheAsnAlaLeuSerValGluGlyGluThrAsnPh 1686
Qy 4291 GGATCAGAAATACGGAG-----ACAGAGCGCGTGGAGAGTATTGCTCTCCAGA 4341
Db 1686 eAsnLeuAsnGlySerSerLeuLeuSerPheAsnGlyAsnSerValPheAsnAlaPro-- 1705
Qy 4342 TGATTCTTAAGTTTGAAGGCAATAAAGGTTCTTATGTTTGTGTTGATTACAAC 4393
Db 1706 -----ValAsnPheTyrAlaAsnAsnSerGlnIleSerPheThrHisSer 1720

RESULT 24

US-11-052-554A-16
; Sequence 16, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16
; LENGTH: 1250
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-16

Alignment Scores:
Pred. No.: 1,35e-12 Length: 1250
Score: 245.00 Matches: 273
Percent Similarity: 31.1% Conservative: 144
Best Local Similarity: 20.3% Mismatches: 385
Query Match: 3.1% Indels: 540
DB: 11 Gaps: 65

US-10-701-844-1 (1-4435) x US-11-052-554A-16 (1-1250)

Qy 553 ACTGTTTTTCTGCAGAGAGTTAAACATTAATAAATCTTGACAACTTATTGACGCTTG 612
Db 20 SerLeuPheSerAla-----AsnGlyValAlaAlaAla 30
Qy 613 CCTTTAAGTTGTTTGGAACTTATTAGGAGTGTCTGTTTGGAGGAGGACACTCG 672
Db 31 IleAspLeuCysGlnGly-----TyrAspIleLysAlaSerCysHisAla 45
Qy 673 TTGACTTTTCGAGAACATACCGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGTCT 732
Db 46 -----SerArgGlnSerLeuSerGlyIleThrGlnValTrpSerIleAla 60
Qy 733 GATGAGCTGTTACTATTGAGGTTTTTAAGAATTATCCTTTTCCAATTGCAATTCTATTA 792
Db 61 AspGlyGlnTrp-----LeuValPheSerAsp----- 69
Qy 793 CTTGCGCTACTGCTGCTGCAACGACTAATAAGGGTAGC----- 831
Db 70 -----MetThrAsnAsnAlaSerGlyGlyAlaValPheLeuGln 82
Qy 832 CAGACTCCGACGACACATCTACACCTCTAATGGTACTATTATTCTTAAACAGATCTT 891
Db 83 GlnGlyAlaGluPheThrLeuSerProGluAsn-----GluThrGlyMet 97
Qy 892 TTGTTACTCAATAATAGAGAGTTCTCATCTATAGTAATTAGTCTCTGGAGAT----- 945
Db 98 ThrLeu-----PheAlaAsnAsnThrValSerGlyGlyTrpAsn 110
Qy 946 ---GGGAGGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTC 1002
Db 111 AsnGlyGlyAlaIlePheAlaLys----- 118
Qy 1003 TTCCAAGAAATATCTGCTCAAGCTGATGGGGAGCTGTGTCAGTAGTAGTACCAGTTTCTCT 1062
Db 119 -----GluAsnSerThr----- 122
Qy 1063 GCTATGGCTAACGAGGCTCCTATTGCCCTTTGCTAGCAATGTTGACAGGAGTAAGAGGGGA 1122
Db 123 -----LeuAsnLeuThrAspValIlePheSerGlyAsnValAlaGlyTyrGlyGly 140

```
QY 1123 GGGATTGCTGCTTCCAGGAGGGCAGCGAGGTGCATCATCTACTTCAACAGAGAT 1182
Db |||||
141 AlaIle-----TyrSerGlyThrAsnAsp 149
QY 1183 CCAGTAGTAAGTTTTCCAGAAATATCGCGTAGAGTTTGTAGGGAACGTAGCCCGA--- 1239
Db |||||
150 ThrGlyAlaIleAspLeuArgValThrAsnAlaValPheArgAsnAsnIleAlaAsnAsp 169
QY 1240 ---GTAGGAGGAGGATTACTCCCTACGGGAACGTTGCTTCTCTGAATATGAAAAACC 1296
Db |||||
170 GlyIleGlyAlaIleThrIleAsnAsnAspIleThrLeuSerAsp-----Asp 187
QY 1297 TTGTTTCTCAACAATGTTCTCTCTGTTTACATGCTGCTTAAGCAACCAACAGTGA 1356
Db |||||
188 ValPheAsnAsnGlnAla----- 194
QY 1357 CAGGCTTCTAATACGAGTAATAATTACGAGATGGAGAGCTACTTCTGTAAGAATGCT 1416
Db |||||
195 ---TyrThrSerThrSerTyrSerAspGlyAspGlyAlaIle----- 208
QY 1417 GCGAAGCAGGATCCATATCTCGGATCAGTTTCTTGTGATGGAGGAGTAGTT--- 1473
Db |||||
209 ---AspValThrAspAsnAsnSerAspSerIleHisProSerGlyThrIleIleAsn 227
QY 1474 ---TTCTTTAGTAGCAATGCTGCTCGGGAAGGGGAGCTATTTATGCCAAAAG 1527
Db |||||
228 AsnThrAlaPheThrAsnAsnThrAlaGluGlyTyrGlyAlaIleIleTyr----- 244
QY 1528 CTCCTGGTCTCAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGAT----- 1581
Db |||||
245 ---ThrAsnSerAlaThrAlaProTyrLeuIleAspIleSerValAspAspSer 261
QY 1581 ----- 1581
Db 262 TyrSerGlnAsnGlyValLeuValAspGluAsnAsnSerAlaAlaGlyTyrGlyAsp 281
QY 1582 -----GTTGGAGCATTTATTTAGGAGATCTGAGAGCTCAGTTTA 1623
Db |||||
282 GlyProSerSerAlaAlaGlyPheMetTyrLeuGlyLeuSer---GluValThrPhe 300
QY 1624 TCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAGACAGCAAGAGAAAT 1683
Db |||||
301 AspIleAlaAspGlyIleThrLeuValIleGlyAsn-----ThrGluAsnAspGly 317
QY 1684 GTCGCGCATTTAATGCGCTAAGTGTCTCTCAAGCCATTTCCGATGGAGTGGGAGGG 1743
Db |||||
318 AlaValAsp-----SerIleAlaGlyThrGly 326
QY 1744 AATATACGACATTAAGACTAAGCAGGCGCATCAGATTCTC----- 1785
Db |||||
327 LeuIleThr-----LysThrGlySerGlyAspLeuValLeuAsnAlaAspAsnAsnAsp 344
QY 1786 TTTAATGATCCCATCAGATGCGCAACGGAATAACACGAGCAGCGAGTCTTCCAAACTT 1845
Db |||||
345 PheThrGlyGluMetGlnIleGluAsnGlyGluValThrLeuGlyArgSerAsnSerLeu 364
QY 1846 CTAATAAATTAAACGATGGTGAA-----GGATACACAGGG 1878
Db |||||
365 MetAsnValGlyAspThrHisCysGlnAspAspProGlnAspCysTyrGlyLeuThr--- 383
QY 1879 GATATTGTTTCTTAATGGAACGAGTACTTTTGTACCAAAATGTTAGTAGAGCAGGA 1938
Db |||||
384 -----IleGlySerIleAspLysTyrGlnAsn----- 392
QY 1939 AGGATTGTTCTCGTGAAGAGGCAAAATATCAGTGAATCTCTAAGTCAGACA----- 1992
Db |||||
393 -----GlnAlaGluLeuAsnValGlySerThrGlnGlnThrPheAla 406
QY 1993 -----GGTGGAGTCTCTATATATGGAAGCTGGGAGTACATGGGAT 2031
Db |||||
407 HisSerLeuThrGlyPheGlnAsnGlyThrLeuAsnIleAspAlaGly----- 422
```



```
Db      637 ValAspAlaGly-----:::-----ValAspThr 643
Qy      3124 CCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATG 3183
Db      644 GlnTrpGlyAlaLeuMetAlaAspSerGlyGlnHisGlnAspGluGlySerThrLeu 663
Qy      3184 GCGGCTTATATCTGTGATGCTTATGCG---ACCATCTCTGTTACTGAGACAAACGCTCTCA 3240
Db      664 ThrLysThrGlyAlaGlyThrLeuGluLeuThrAlaSerGlyThrGlnSerAlaVal 683
Qy      3241 TCCCATCAAGACACATGG-----ACAACAGATGCTTTTCATTATGACAGACAT 3288
Db      684 ArgValGluGluGlyThrLeuGlnGlyAspValAlaAspIlePheProTyrAlaSerSer 703
Qy      3289 GGAGTTGTGTTAGAGGATCTATGATGCTTCTCTAACA-----AGTAATATA 3336
Db      704 LeuTrpValGlyAspGlyAlaThrPheValThrGlyAlaAspGlnAspIleGlnSerIle 723
Qy      3337 GAAGTATATGCCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTGATG 3396
Db      724 AspAlaThrSerSerGlyThrIleAspIleSerAspGlyThr-----737
Qy      3397 GCAGGAATGAGTCCGGTCTTAAATAATGTTAGATAGTTAAGTTAGCGATGCT 3456
Db      738 -----ValLeuArg 740
Qy      3457 TTTTCTTGAGATCATCATCTATTTTGTGTTTGTAGCTGTTG-----TGTCTCTATTTCG 3510
Db      741 LeuThrGlyGlnAspThrSerValAlaLeuAsnAlaSerLeuPheAsnCysAspGlyThr 760
Qy      3511 TATGATTCGGAGTCTCTCTCAAGTGTATAGCCTTAATGTAACCACTCTCTTTTAAAGGA 3570
Db      761 LeuValAsnAlaThrAspGlyValThrLeuThrGlyGluLeuAsnThrAsnLeuGluThr 780
Qy      3571 GACGATGTT---TACTGTAATGGAGACTGCGGCTTTGTCTAATGTTCTATGACGAGCTGAA 3627
Db      781 AspSerLeuThrTyrLeuSer-----787
Qy      3628 GAAGTTTCGATTATCTCAGCTAATGGCGACAATTTAAGCATTTACCGGACAAAACCATACA 3687
Db      788 -----AsnValThrValAsnGly-----793
Qy      3688 TTATCATTTACAGATTCCTCAAGGCGCAGTT---CTTCAAAATATATGCTCTCATTTTCAGA 3744
Db      794 ---AsnLeuThrAsnThrSerGlyAlaValSerLeuGlnAsnGlyVal-----Ala 809
Qy      3745 GGAGAGACATTTACTCTCAGA---GATTTTTCG-----AGTCTGATGTTCTCGAAA 3792
Db      810 GlyAspThrLeuThrValAsnGlyAspTyrThrGlyGlyThrLeuLeuLeuAspSer 829
Qy      3793 AATGTT-----TCTTGGCGAGAAAAGGAATGATCTCGGGAACCCGTG---3837
Db      830 GluLeuAsnGlyAspAspSerValSerAspGlnLeuValMetAsnGlyAsnThrAlaGly 849
Qy      3838 -----AGTATTTCCGAGCAGCGGAA-----3858
Db      850 AsnThrThrValValAsnSerIleThrGlyIleGlyGluProThrSerThrGlyIle 869
Qy      3858 -----3858
Db      870 LysValValAspPheAlaAspProThrGlnPheGlnAsnAsnAlaGlnPheSerLeu 889
Qy      3859 -----GTGATTTTCTGGGATACTCC 3879
Db      890 AlaGlySerGlyTyrValAsnMetGlyAlaTyrAspTyrThrLeuValGluAspAsnAsn 909
Qy      3880 GTGGGGTAT-----TCTCTTTTATCTCTGTCGCAACC 3912
Db      910 AspTyrTyrLeuArgSerGlnGluValThrProProSerProAspProAspProThr 929
Qy      3913 TCATCATCAACTCCG-----CCTGCTCCAACAGTTAGTATGCTCGGAAAGGGTCT 3963
```

```
Db      930 ProAspProAspProThrGlnAspProAspProThr---ProAspProGluProThrPro 948
Qy      3964 ATTTTTCCTAGAGACTAGTTTCGAGATCTCAGGCGCTCAAAAAGGGTCTATGTTTCAT 4023
Db      949 AlaTyrGlnProValLeuAsnAlaLysValGlyGly-----TyrLeu 962
Qy      4024 AATAATCCCGGAATTTCCGACACAGTTTTCGAGGTAAAGTAATAATAATGCTGGTGGT 4083
Db      963 AsnAsnLeuArgAlaAlaAsnGlnAlaPheMetMetGluArgArgAspHisAlaGlyGly 982
Qy      4084 GGAGGC 4089
Db      983 AspGly 984
RESULT 25
US-11-052-554A-172
; Sequence 172, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 172
; LENGTH: 1643
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-11-052-554A-172
Alignment Scores:
Pred. No.: 2,8e-12 Length: 1643
Score: 242.00 Matches: 282
Percent Similarity: 33.2% Conservative: 188
Best Local Similarity: 19.9% Mismatches: 526
Query Match: 3.1% Indels: 419
DB: 11 Gaps: 70
US-10-701-844-1 (1-4435) x US-11-052-554A-172 (1-1643)
Qy      604 GCAGCTTTCCTTAAAGTTGTTTGGGAACCTTATTAGGAGTGTCTTACTGTTTATGGGAGA 663
Db      14 AlaGlyLeuValThrAlaSerThrAlaThrIleValAlaGlyPheSerGlyValAlaMet 33
Qy      664 GCAGCTCGTTGACTTTTCGAGAACATACGGACTTCTACAAATGGCGGAGCTCTAAGTAAT 723
Db      34 GlyAlaAlaMetGlnTyrAsnArg-----ThrThrAsnAlaAlaAla-----47
Qy      724 AGCGTCTGATGAGTCTGTTTACTATTGAGGGTTTAAAGATTAATCTCTTTCGAATTC 783
Db      48 ThrThrPheAspGlyIle-----GlyPheAspGlnAlaAlaGlyValaAsn---62
Qy      784 AATTCATTACTTTCGCTGCTGCTGCAACGACTAATAAGGGTAGCAGACTCCGACG 843
Db      63 -----IleProValAlaProAsnSerValIleThrAlaAsnAlaAsnProIle 79
Qy      844 ACAACATCTACACCGCTCTAATGGTACTATTATTCT-----AAAAACAGAT 888
Db      80 ThrPheAsnThrPro---AsnGlyHisLeuAsnSerLeuPheLeuAspThrAlaAsnAsp 98
Qy      889 CTTTGTGTTACTCAATAATGAG-----909
Db      99 LeuAlaValThrIleAsnGluAspThrThrLeuGlyPheIleThrAsnIleAlaGlnGln 118
Qy      910 ---AAGTCTCTCATCTTATAGTAATTTAGTCTCTCTGAGATGGGGAGCTATAGATGCTAAG 966
```

119 AlaLysPhe---PheAsnPheThrValAlaAlaGlyLysIleLeuAsnIleThrGlyGln 137
967 AGCTTAACGGTTCAAGGAATTAAGCAAGCTTTGTGCTTCCACAGAAATACTGCTCAAGCT 1026
138 GlyIleThrValGlnGlnAlaSerAsnThrIleAsnAlaGlnAsnAlaLeuThrLysVal 157
1027 GATGGGGAGCTTGTCAAGTAGTCACCAAGTTCTCTGCTATGCTTAAC----- 1074
158 HisGlyGlyAlaAlaIleAsnAlaAsnAspLeuSerGlyLeuGlySerIleThrPheAla 177
1075 -----GAGGCTCTATT 1086
178 AlaAlaProSerValLeuGluPheAsnLeuIleAsnProThrThrGlnGluAlaProLeu 197
1087 GCCTTTGTAGCAATGTTCCAGAGTAAGAGGGGA---GGGATTGCTGCTGTTCCAGAT 1143
198 ThrLeuGlyAlaAsnSerLysIleValAsnGlyGlyAsnGlyThrLeuAsnIleThrAsn 217
1144 GGGCAGCAGGGAGTGCATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCACGA 1203
218 GlyPheIleGlnValSerAspAsnThrPheAlaGlyLysIleThrIleAsnIleAspAsp 237
1204 AATACTCGCGTAGAGTTGTATGGGAACGTA-----GCCGAGTA 1242
238 CysGlnGlyLeuMetPheAsnSerThrProAspAlaAlaAsnThrLeuAsnLeuGlnVal 257
1243 GGAGGA-----GGGATTACTCTACGGGAACGTTGTTTCCTG--- 1281
258 GlyGlyAsnThrIleAsnPheAsnGlyIleAspGlyThrGlyLysLeuValLeuValSer 277
1282 ---AATAATGGAAACCTTGTT----- 1302
278 LysAsnGlyAlaAlaThrGluPheAsnValThrGlyThrLeuGlyAsnLeuLysGly 297
1303 -----CTCAACAATGTTGCTTCTCTCTGTTTACATGCTGCTAAGCAACCAACAAAGT 1353
298 IleIleGluLeuAsnThrAlaAla-----ValAlaGlyLysLeuIleSerGln 313
1354 GGACAGCTTCTAAT-----ACGAGTAATAATTCAGGAGAT----- 1389
314 GlyGlyAlaAlaAsnAlaValIleGlyThrAspAsnGlyAlaGlyArgAlaAlaGlyPhe 333
1390 -----GGAGGAGTATCTTCTGTAAGAAT 1413
334 IleValSerValAspAsnGlyAsnAlaAlaThrIleSerGlyGlnValThrAlaLysAsn 353
1414 GGTGGCAACAGGATCCAACTCTGATCACTGCTGATCACTTCTGATGGAGGAGGAGTAGTT 1473
354 MetValIleGlnSerAlaAsnAlaGlyGlyGlnValThrPheGlu----- 368
1474 TTCTTTAGTAGCAATGTAGTGTGGGAAAGGGGAGCTATTATTATGCGCAA----- 1524
369 -----HisIleValAspValGlyLeuGlyGlyThrAsnPheLysThrAlaAsp 385
1525 ---AAGCTCGGTT-----GCTAACTGGGCGCTGTACAAATTT----- 1560
386 SerLysValIleThrGluAsnSerAsnPheGlySerThrAsnPheGlyAsnLeuAsp 405
1561 -----TTAAGG 1566
406 ThrGlnIleValValProAspThrLysIleLeuLysGlyAsnPheIleGlyAspValLys 425
1567 AATATCGCTAATGATGTGGAGCGGATTTATTAGGAGAATCTGGAGAGCTC---AGTTTA 1623
426 AsnAsnGlyAsnThrAlaGlyValIleThrPheAsnAlaAsnGlyAlaLeuValSerAla 445
1624 TCTCTGATTAGGAGATATTATTTTCGATGGGAATCTTAAAGAACACAGCCCAAGAGAT 1693
446 SerThrAspProAsnIleAlaValThrAsnIleAsnAlaIleGluAlaGluGlyAlaGly 465
1684 GCTCCCGATGTTATGGGTAACTGTGCTCTCACAGCCATTTTCGATGGGATCGGAGGG 1743
466 ValValGluLeuSerGlyIleHisIleAlaGlu-----LeuArgLeuGlyAsnGlyGly 483

1744 AAAATAACGACATTAAAGAGCTAAAGCAGGCGATCAGATTCTTTTAATGATCCCATCGAG 1803
484 SerIle-----PheLysLeuAlaAspGlyThrValIleAsnGlyProValAsn 499
1804 -----ATGCGAAACGGAATTAACAGCAGCGAGCTTCCAAACTTCTTAAAAATT 1854
500 GlnAsnAlaLeuMetAsnAsnAsnAlaLeuAlaGlySerIleGlnLeu----- 516
1855 AACGATGTGAAGGA---TACACAGGGGATATTGTTTTGCTTAATGGAAGC---AGTACT 1908
517 ---AspGlySerAlaIleIleThrGlyAspIle-----GlyAsnGlyGlyValAsnAla 533
1909 TTGTACCAAAATGTTACGATA-----GAGCAAGGAAGGATGTTCTTCGTAAGGAAGCA 1962
534 AlaLeuGlnHisIleThrLeuAlaAsnAspAlaSerLysIleLeuAlaLeuAspGlyAla 553
1963 AAATATTACGTGAATCTTAAGTCAGACAGTGGAGTCTGTATATGGAAGCT---GGG 2019
554 AsnIleIleGly-----AlaAsnValGlyGlyAlaIleHisPheGlnAlaAsnGly 570
2020 ACTACATCGGATTTGTAATCCACCAACCAACAGCCTCTGCGCTAATCAGTTG 2079
571 GlyThrIleLysLeuThrAsnThrGln-----AsnAsnIle 582
2080 ATCAGCTTTTCCAATTCGCAATTTGCT-----CTTCTTCT 2115
583 ValValAsnPheAspLeuAspIleThrThrAspLysThrGlyValValAspAlaSerSer 602
2116 TTGTAGCAAAACATGAGTACGATACGATCTCTCTACCAATCTCCAGCGCAAGATTCTCAT 2175
603 LeuThrAsnAsnGlnThrLeuThrIleAsnGlySer----- 614
2176 CTGCAAGTCTTGTAGCACAACTGCTGCTTCTGTACAAATAGTAGGCGCTATCTTTT 2235
615 -----IleGlyThrValValAlaAsnThrLysThrLeuAla 626
2236 GAGGATTGCGATGATACAGCTTATGATAGGTATGATTGGCTAGTGTCTAATCAAAAATC 2295
627 ---GlnLeuAsn-----IleGlySerSerLysThrIle 636
2296 AATGCTCTGAAATTAACAGTTAGGAGCTAAGCCCGCAGCTAATGCCCCATCAGATTGACT 2355
637 -----LeuAsnAlaGlyAspVal-----AlaIleAsnGluLeuVal 648
2356 CTAGGAATGAGATGCTAAGTATGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCCT 2415
649 IleGluAsnAsnGlySerValGlnLeuAsnHisAsnThrTyrLeuIleThrLysThrIle 668
2416 AATACAGCAAAATAATGCTCTTATCTCTGAAAGCTTACATGGAATAAACTGGGTATAAT 2475
669 AsnAlaAlaAsnGlnGlnIleIleValAlaAlaAspProLeuAsnThrAsn----- 686
2476 CTGGGCGCTGAGCAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCATTATAGAT 2535
687 -----ThrThrLeuAlaAspGlyThrAsnLeuGlySerAlaGluAsn 700
2536 ---ATGAGATCTCGGATTCAGCAATTCAGCA-----AGTGGGATGGCGC 2580
701 ProLeuSerThrIleHisPheAlaThrLysAlaAlaAsnAlaAspSerIleLeuAsnVal 720
2581 TCTATTGTCGAGGATTTGTTCTGGAGTTTTCGAATTTCTTCTATCATGACCGCGAT 2640
721 GlyLysGlyValAsnLeuThrAlaAsnAsnIleThrThr-----AsnAsp 735
2641 GCCTTAGTTCAGGATATCGGTATATATAGTGGGGTTATTCTTTAGGAGCAAACTCTAC 2700
736 AlaAsnValGlySerLeuHisPheArgSerGlyGlyThrSerIleValSerGlyThrVal 755
2701 TTTGGATCATCGATGTTGTTGCTAGCATTTTACCGAAGTATTTCGTAGATCTAAGATTAT 2760
756 GlyGlyGlnGln-----GlyHisLysLeuAsnAsn 765

```

QY 2761 GTAGTGTGCTTCCATCATCATCGTTCGATAGATCGGT-----TATCTATCTACC 2814
Db      :::::  :::::  |||:::  |||:::  :::::  :::::  :::::  :::::  :::::
766 LeuileLeuAaspAsn-----GlyThrThrValValPheLeuGlyAasp 779
QY 2815 CAACAAGCTTTATGTGGATCCCTATTGTTTGGAGATGCGTTTATCCGTGCTAGCTACGGG 2874
Db      |||:::  :::::  :::::  |||:::  |||:::  |||:::  |||:::  |||:::
780 ThrThrPheAenGlyGlyThrlyleleGluGlyLysSerleleLeuGlnleleSerAenAen 799
QY 2875 TTTGGGAATCAGCATATCAAAACCTCATATACATTTGCAGAGGAGAGCGATGTTCTGTGG 2934
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
800 TyrThrThrAaspHieValGluSerAlaAaspAenThrGlyThrLeuGluPheValAenThr 819
QY 2935 GATAATACTGTCTGGCTGGAGAGATTGGAGCGGATTAACGATGTGATTAATCCATCT 2994
Db      |||  |||  |||  |||  |||  |||  |||  |||
820 Asp-----ProileThrValThrLeuAen 827
QY 2995 AAG-----CTCTATTGTAATGAGTTCGCTCTTTCGTGCAAGCTGAGTTCCTTATGCC 3048
Db      |||  |||  |||  |||  |||  |||  |||  |||
828 LysGlnGlyAlaTyPheGlyValLeuLys-----GlnValleleleSerGlyPro 844
QY 3049 GATCATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGGCATTCAGAGCGGACATCTC 3108
Db      :::::  |||  |||  |||  |||  |||  |||  |||
845 GlyAenlleValPheAenGluileGly-----853
QY 3109 CTAATCTATCATGCTCTGTTGGAGTGAAGTTTGATGATGTTCTAGTACACATCTTAAT 3168
Db      |||:::  :::::  |||:::  |||:::  |||:::  |||:::  |||:::
854 ---AenValGlyIleValHieGlyIleAla-----AlaAen 864
QY 3169 AAATATAGCTTTATGCGCGCTTATCTGTGATGCTTATCGCACCATCTCTGGTACTGAG 3228
Db      |||  |||  |||  |||  |||  |||  |||  |||
865 SerleleSerPheGluAenAlaSerLeuGlyThrSerLeuPheLeuProSerGlyThrPro 884
QY 3229 ACAACGCTCTATCCATCAAGAGACA-----TGGACAACAGATGCCCTTTCATTTA 3279
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
885 LeuAaspValLeuThrIleLysSerThrValGlyAenGlyThrValAaspAenPheAenAla 904
QY 3280 GCACACATGAGTTGTTGTTAGAGATCTATGATGCTTCTTAACAGTAATATAGAA 3339
Db      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
905 Profile-----ValValValSerGly-----IleAaspSerMetIleAenAenGlyGln 920
QY 3340 GTATATGCGCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTTGAGTGCA 3399
Db      :::::  |||  |||  |||  |||  |||  |||  |||
921 llelleGly-----AapLysLysAenlleleleleAenSerLeu 933
QY 3400 GGAAGTAGATCCGGTCTTAAATAATTTGGTTAGATAGTTAAGTGTAGCGATGCTTTT 3459
Db      |||  |||  |||  |||  |||  |||  |||
934 Gly-----934
QY 3460 TCTTTGAGATCATCATCATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3519
Db      |||  |||  |||  |||  |||  |||  |||
934 -----934
QY 3520 GCGAGCTCTCCTCAAGTGTAAAGCCCTAATGTAACCACTCTCTTTTAAAGGAGACGATGT 3579
Db      |||  |||  |||  |||  |||  |||  |||
935 -----SerAaspAenSerleleThrValAenAlaAenThrLeutySerGlyleAargThr 952
QY 3580 TACTTGAATGGAGACTGCGCTTTTGTCT-----AAT 3609
Db      |||  |||  |||  |||  |||  |||  |||
953 ThrLysAenAenGlnGlyThrValThrLeuSerGlyGlyMetProAenAenProGlyThr 972
QY 3610 GTCTATGCA---GGAGCTGAAGAAGGTGCGATTATCTAGCTAATGCGGACAAATTAACG 3666
Db      :::::  |||  |||  |||  |||  |||  |||
973 lleTyGlyLeuGlyLeuGluAenGly-----SerProLysLeuLysleGlnValThr 989
QY 3667 ATTACCGGCAAAACCATATCATTTACAGATTTCTCAAGGCCAGTTCCTCAAAAT 3726
Db      :::::  |||  |||  |||  |||  |||  |||
990 PheThrThrAaspTyAenAenLeu-----GlySerleleleleAen 1003
QY 3727 TATGCTTTTCAATTCAGCAGGAGACACTTACTCTCAGAGATTTTTCGAGTCTGATGTC 3786
Db      |||:::  |||:::  |||:::  |||:::  |||:::  |||:::  |||:::
1004 AenValThrIleAenAaspTyValThrLeuThrThrGlyGlyIleAleGlyThrAaspPhe 1023
QY 3787 TCGAAAAATGTTTCTTGGCGAGAAAAGGAATGATCTCGGGAACACCGTGATATTCC 3846

```

```

Db 1024 AspAlaLysIleThrLeuGly-----|||:::  |||:::  :::::  :::::  :::::  1033
QY 3847 GGAGCAGCGCAAGTGATTTTCTGGGATAACTCCGTCGGGTATTCTCTTTTATCTACTGTG 3906
Db      |||  |||  |||  |||  |||  |||  |||  |||
1034 GlyAenAlaAenValAargPheValAaspSerThrPheSer---AaspProArgSerMetIle 1052
QY 3907 CCAACCTCATCATCAACTCCGCTGCTCCACAGCTTAGTGTCTCGGAAGGGTCTATT 3966
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
1053 -----ValAlaThrGlnAlaAenLysGlyThrVal 1062
QY 3967 TTTTCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAGGGGTCTATGTTCTGATAAT 4026
Db      |||  |||  |||  |||  |||  |||  |||  |||
1063 ThrTyrLeuGlyAenAlaLeuVal-----Ser 1071
QY 4027 AATCGCGGAATTTTCGGAACAGTT-----TTTCGAGGTGAAGATAATAAT 4071
Db      |||  |||  |||  |||  |||  |||  |||  |||
1072 AenlleGlySerLeuAaspThrProValAlaSerValargPheThrGly-----Aen 1088
QY 4072 AATGCTGTGTGTGGAGCAGTGGGTTCGCTACACCATCAAGTAGCAGCTTTTACAGTTAA 4131
Db      :::::  |||  |||  |||  |||  |||  |||  |||
1089 AspSerGlyAlaGlyLeuGlnGlyAenlleTySerGlnAenlleAaspPheGlyThrTy 1108
QY 4132 AAACGTGTAACGGAAAGTTTCTTTCACAGATAAGCTAGCTCTTCGCGAGCGGAGTGT 4191
Db      |||  |||  |||  |||  |||  |||  |||  |||
1109 AenLeu-----ThrIleLeuAen-SerAenValIleLeuGlyGlyThrAl 1125
QY 4192 TTATAAAGGC---ATTGTGCTTTTCAAGACACATGAAGGAGGCATATTCTTCCGAGGAA 4248
Db      |||  |||  |||  |||  |||  |||  |||  |||
1125 alleaenGlyIleAaspLeuLeuThrAenAenLeu-----IlePheAlaAenGlyTh 1143
QY 4249 CACAGCATACATGATTTAAGGATTTCTGTGCTACTAATCAGGATCAGAATACGAGAC 4308
Db      :::::  |||  |||  |||  |||  |||  |||  |||
1143 rSerThrTrpGlyAaspAenThrSerIleSerThrThr-----LeuAenValSerSe 1160
QY 4309 AGGAGCGGTGGAGGATTTTGTCTCTCCAGATGATTTCTCTGA 4351
Db      |||  |||  |||  |||  |||  |||  |||  |||
1160 rGlyAenlleGlyGlnValValIleAlaGluAaspAlaGlnVal 1174

RESULT 26
US-11-052-554A-281
; Sequence 281, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 281
; LENGTH: 5291
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-281

Alignment Scores:
Pred. No.: 1,76e-11 Length: 5291
Score: 235.50 Matches: 287
Percent Similarity: 34.2% Conservative: 193
Best Local Similarity: 20.5% Mismatches: 537
Query Match: 3.0% Indels: 386
DB: 11 Gaps: 63

US-10-701-844-1 (1-4435) x US-11-052-554A-281 (1-5291)
QY 496 GGGGAGCGTTAACTGTA-----TCATTTCCTTACTACTGTTATAGAGAT 540

```


QY 2490 ----- 2490
Db 3934 AlaProAlaMetGlyIleAsnIleAspSerLeuGlnAlaAepThrGlyLeuSerAlaSer 3953
QY 2491 -----GTAGCTCTTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATACGATCT 2544
Db 3954 AspPheIleThrSerValSerProValValValAsnGlySerLeuThrAlaAlaLeuAla 3973
QY 2545 CGCATTCAGCAAAATCAAGCAAGTGTGATGGCGCTCTTATTTGCGAGGATTTAGGTT 2604
Db 3974 SerAsnGluThrAlaGlnIleSerIleAspGlyGlyThrThrTrpThrLeuThrVal 3993
QY 2605 TCTGAGTTTCGAATTTCTTATCATACACCGCGATGCTTTAGTTCAGGGA----- 2655
Db 3994 ThrGly-----ThrThrTrpArgTyrAsnAspSerArgThrLeuThrAspGlyAsnTyrLeu 4012
QY 2656 -----TATCGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCCTACTTTGGATCA 2709
Db 4013 TyrGlnValArgValIleAspAlaAlaGlyAsnValGlyAlaThrAsp----- 4028
QY 2710 TCGATGTTTGGTCTAGCATTTACCGAAGTATTGGTAGATCTAAAGATTATGTAGTGTG 2769
Db 4029 -----SerGlnAsnValValIleAsp 4035
QY 2770 GGTTCATATCATGCTTGCATAGATCCGTTTATCTATCTACC-----CAA 2817
Db 4036 ThrThrAlaProAspProAlaValIleAspValIleSerAlaIleThrThrAspMet 4055
QY 2818 CAAGCTTTATGGATCTCTTTGTTGGAGATGGTTTATCCGCTAGCTACGGGTTT 2877
Db 4056 GlyLeuIleThrAsnAspPheValThrSerAspThrThrLeuAlaValSerGlyThrLeu 4075
QY 2878 GGGATCAGCATATCAAACTCATATACATTTGCGAG-----GAGAGCGATGTT 2928
Db 4076 Gly-----AlaThrLeuSerAlaGlyGluPheAlaGlnIleSerLeuAspGlyVal 4093
QY 2929 COTGGGATAAATCTGCTGGCTGGAGAGATTGGAGCGGGATTACCGATTGTGATTACT 2988
Db 4094 ThrTrpThrLeuThrValValGly----- 4102
QY 2989 CCATCTAAGCTCTATTGTAAGTAGTTGGCTCTTTCGTCGCAAGCTGAGTTTCTTATGCC 3048
Db 4103 -----ThrSerTrpSerTyrAla 4108
QY 3049 GATCATGAATCTTTACAG-----GAA 3072
Db 4109 AspGlyHisThrLeuThrAspGlyThrTrpAsnTyrThrValArgValValAspLeuAla 4128
QY 3073 GCGCATCAAGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTTTGGA 3132
Db 4129 GlyAsnValGlyGlnThrAlaThrGlnAsnValValAspThrThrSerProGluAla 4148
QY 3133 GTGAGTTTGATCGATGTTCTAGTACACATCTTAATAATAGCTTTATGCGGCTTAT 3192
Db 4149 AlaLys----- 4150
QY 3193 ATCTGTGATCTTATCGCAACCATCTCTGGT-----ACTGAGACAACGCTCTATCCCATCAA 3249
Db 4151 -----SerIleThrIleThrGlyIleSerAspAspThrGlyThrSerSer 4166
QY 3250 GAGACATGGAACACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTGTTAGAGATCT 3309
Db 4167 AspPheIleThrSerAspThr-----ThrLeuThrValArgGlyVal 4180
QY 3310 ATGTATGCTTCTTAAAGTAAATATAGAGTATATGCGCATGGAGATATAGTATCGA 3369
Db 4181 LeuGlyAlaAlaLeuGlyAlaAsn-----GluPheAla 4191
QY 3370 GATGCTTCTCGAGGCTATGTTGTTAGTGCAGGAAGTAGAGTCGGGTTCTAAAAAATATCG 3429
Db 4192 GlnIleSerThrAspAsnGlyAlaThr-----Trp 4201

QY 3430 TTAGATAGTTAAGTGTAGCGATGCCCTTTTCTTTTGGAGATCTACATCATTTTGTTTT 3489
Db 4202 ValAsnValThrValAlaAlaAepSerLeuAsn----- 4212
QY 3490 GCTTGTGTGTTCTTATTCGTATGATTCGCGAGTCTCTCTCAAGTGTTTAAACGCTAAT 3549
Db 4213 -----TrpSerTyrValAspGly-----ArgThrLeuThr-----Asn 4223
QY 3550 GTAAACCATCTCTTTTAAAGGAGACGATGTTTACTTGAATGGAGACTCGCGCTTTTGTCAAT 3609
Db 4224 GlyThrThrThrTrpGlnValValArgValValAspLeuAlaGly-----Asn 4238
QY 3610 GTCTATCGAGAGCTGAAGAAGTTCGATTTATCTCAGCTAATGGCACAATTTAACGATT 3669
Db 4239 ValGlyAlaThrSerSerGlnSerAlaLeuIle----- 4249
QY 3670 ACCGACAAAACCATACATATTATTCATATTTCACAGATCTCAAGGGCCAGTT-----CTTCAAAAT 3726
Db 4250 -----AspThrValAsnProAlaGlnValLeuThr 4259
QY 3727 TATGCTTTCATTTTCCAGGAGAGACACTTACTCTGAGAGATTTT-----TCGAGTCTG 3780
Db 4260 IleAlaSerIleSerThrAspThrGlySerSerAlaThrAspPheIleThrSerAspThr 4279
QY 3781 ATGTTCTCGAAAAATGTTTCTTGGGAGAAAGGAATGATCTCCGGAAAACCGTGAGT 3840
Db 4280 MetLeuThrLeuThrGlySerLeuGly-----AlaGlyLeuAlaSerGlyGluValAlaGln 4298
QY 3841 ATTTCCGAGCAGCGGAAGTATTTCTGCGATTAATCTCCGTGGGTATTTCTCTTTATCT 3900
Db 4299 IleSer-----LeuAspSerGlyAlaThrTrpThrLeuThr 4311
QY 3901 ACTGTGCAACCTCATCATCACTCCGCTGCTCCAAAGTAGTAGTGTGCGGAAAGG 3960
Db 4312 ThrAsnGlyThrGlnTrpThrTyrThrAspSerArgThrLeuThrAsp-----Gly 4328
QY 3961 TCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAGGGTCTATGTTT 4020
Db 4329 SerTyrValTyrGlnValArgValLeuAspLeu----- 4339
QY 4021 GATAATAATCGCGGAATTTCCGAAACAGTTTTTCGAGGTAAAGATAATAATATCTGCT 4080
Db 4340 -----AlaGlyAsnThrGlyProValVal-----SerLysThr-ValVal 4352
QY 4081 GGTGAGCAGTGGT-----TCCGCTACACCAATCAAGTAGTAGCTTTTACA-----GTAA 4131
Db 4352 lValAspThrIleAsnProThrAlaThrProThrIleValSerTyrThrAspAspValGl 4372
QY 4132 AAACGTGTAAGGGAAGTTTCT-----TTCACAGATAACGTAGCTCT----- 4174
Db 4372 yGlnArgGlnGlyThrLeuSerSerSerGlnAlaThrAspAspThrThrProLeuLeuAs 4392
QY 4175 -----TGGGAGGGGAGTGGTTTAT-----AAAAGCAT 4203
Db 4392 nGlyValLeuSerAlaProLeuAlaSerGlyGluValValTyrLeuTyrArgAsnGlyLe 4412
QY 4204 TGTGCTTTTAAAGCAATGAAGGAGCATATTTCTCCGAGG-----AACACAGATA 4257
Db 4412 uLeuLeu-----GlyAlaValThrMetValGlyAlaLeuAsnTrpThrTy 4427
QY 4258 CGATCAT 4264
Db 4427 rSerAsp 4429

RESULT 27

US-11-052-554A-2
; Sequence 2, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 2
 ; LENGTH: 1571
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli 0157:H7
 US-11-052-554A-2

Alignment Scores:
 Pred. No.: 3,42e-10 Length: 1571
 Score: 218.00 Matches: 218
 Percent Similarity: 32.7% Conservative: 138
 Best Local Similarity: 20.0% Mismatches: 413
 Query Match: 2.8% Indels: 320
 DB: 11 Gaps: 54

US-10-701-844-1 (1-4435) x US-11-052-554A-2 (1-1571)

QY 442 TTAATGGGGGGGATATGCGAGCAAAATCATGTTCTCAAGCAATTTACGATGGGAG 501
 DB :||||| ThrGlnAsnIleAsnAsnTyrGlyIleAlaThrGlyThr 432
 QY 502 ACCTTAATCTATCATTTCCCTATCTATGTTATAGGAGATCCGATGGGACTGTTTTT 561
 DB :||||| ThrGlnAsnSerGlyThrGlnAsnIleLysSerGlyGlyLysAlaAspThrThrIle 452
 QY 562 TCTGAGGAGATTAACTTAAATAATCTTGACAAATCTATTGACGCTTTCCTTAACT 621
 DB :||||| SerSerGlySerArgGlnValGluLysAspGlyThrAlaIleGlySerAsnIleSer 472
 QY 622 TGTTTGGGAATTAATAGGAGTTTACTGTTTATAGGAGAGACACTCGTTGACTTTC 681
 DB :||||| AlaGlyGlySerLeuIleValTyrThrGlyIleAlaHisGly 487
 QY 682 GAGAACTACCGACTTCTACAAATGGGGCAGCTCTTAAGTAATAGCGCTGCTGAGACTG 741
 DB :||||| ValAsnGlnGluThrGlySerAlaLeuValAlaAsnThrGlyAlaGly 503
 QY 742 TTTACTATTGAGGTTTAAAGAAATATCC-----TTTCCAAATTCGAATTCATTA 792
 DB :||||| ThrAspIleGluGlyTyrAsnLysLeuSerHisPheThrIleThrGlyGlyGluAlaAsn 523
 QY 793 CTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
 DB :||||| TyrValValLeu-----GluAsnThrGlyGluLeuThrValValAlaLysThr 539
 QY 853 ACACCGCTAATGCTATTT-----TATTCCTAAA 882
 DB :||||| SerAlaLysAsnThrThrIleAspAlaGlyLysLeuIleValGlnLysGluAlaLys 559
 QY 883 ACAGATCTTTTGTACTCAATAAT-----GAGAAAGTTCTCATTTCTAT 924
 DB :||||| ThrAspSerThrArgLeuAsnAsnGlyGlyValLeuGluValGlnAspGlyGlyGluAla 579
 QY 925 AGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGACTTAACG----- 975
 DB :||||| LysHisValGluGlnGlnSerGlyGlyAlaLeuIleAlaSerThrThrSerGlyThrLeu 599
 QY 976 GTTCAAGCAATTTAGCAAGCTT---TGCTCTTC-----CAAGAAAAT 1014
 DB :||||| IleGluGlyThrAsnSerTyrGlyAspAlaPheTyrIleArgAsnSerGluAlaLysAsn 619
 QY 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCACTTCTCTGCTATG----- 1068
 DB :||||| ValValLeuGluAsnAlaGlySerLeuThrValValThrGlySerArgAlaValAspThr 639
 QY 1069 -----GCTAAC-----GAGGCTCTATTGCTTTGTAGCGAAT 1101

DB 640 IleIleAsnAlaAsnGlyLysMetAspValTyrGlyLysAspValGlyThrValLeuAsn 659
 QY 1102 GTTGCAGG-----GTAAGAGG 1119
 DB 660 SerAlaGlyThrGlnThrIleTyrAlaSerAlaThrSerAspLysAlaIleIleGly 679
 QY 1120 GGA-----GGGATTGCTGCT-----GTTCCAGATGGGACGAGGGA 1155
 DB 680 GlyLysGlnThrValTyrGlyLeuAlaThrGluAlaAsnIleGluSerGlyGluGlnIle 699
 QY 1156 GTGTCAATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATACGCGTA 1215
 DB 700 ValAspGlyGly---SerThrGluLysThrHisIleAsnGlyGlyThrGlnThrValGln 718
 QY 1216 GAGTTT----- 1221
 DB 719 AsnTyrGlyLysAlaIleAsnThrAspIleValSerGlyLeuGlnGlnIleMetAlaAsn 738
 QY 1222 -----GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGAACGTT 1272
 DB 739 GlyThrAlaGluGlySerIleIleAsnGlyGlySerGlnIleValAsnGluGlyGlyLeu 758
 QY 1273 GCT-----TTCCTGAATTAATGGAACCTTGTCTTCTCAACAATGTTGCTTCTCT 1323
 DB 759 AlaGluAsnSerValLeuAsnAspGlyGlyThrLeuAspValArgGluLysGlySerAla 778
 QY 1324 GTTTACATTGCTGCTTAAGCAACCAACAGTGA-----CAGGCTTCT 1365
 DB 779 ThrGlyIle-----GlnGlnSerSerGlnGlyAlaLeuValAlaThrThrArgAlaThr 796
 QY 1366 AATACAGTAATAATATACGAGATGAGGAGCTATCTTCTTAAGAAATGTCGCGCAAGCA 1425
 DB 797 ArgValThrGlyThrArgAlaAspGlyValAlaPheSerIleGluGlnGlyAla----- 814
 QY 1426 GGATCAATAACTCTCGATCAGTTCTTCTTGTAGTAGAGAGGAGTAGTTTCTTTAGTAGC 1485
 DB 815 ---AlaAsnAsn-----IleLeuLeuAlaAsnGlyGlyValLeuThrValGluSer 830
 QY 1486 AATGATAGCTGCTGGGAAA-----GGGGGAGCTATTTATGTCCTCAAAAAGCTC 1530
 DB 831 AspThrSerSerAspLysThrGlnValAsnThrGlyGlyArgGluLeuValLysThrLys 850
 QY 1531 TCGGTTGCTAAC-----TGTGGCCCTGTACAAATTTTAAAGAAATATCGCT--- 1575
 DB 851 AlaThrAlaThrGlyThrThrLeuThrGlyGlyGluGlnIleValGluGlyValAlaAsn 870
 QY 1576 -----AATGATGCTGAGCGATTTATTAGGAGAACTCTGGAGAGCTCAGTTA 1623
 DB 871 GluThrThrIleAsnAspGlyGlyIleGlnThrValSerAlaAsnGlyGluAlaIleLys 890
 QY 1624 TCTGCT-----CATTTATGAGATATTTTTCGATGGGATCTTAAAGAACAA----- 1671
 DB 891 ThrThrIleAsnGluGlyThrLeuThrValAsnAspAsnGlyLysAlaThrAspIle 910
 QY 1672 -----GCCAAAGAGAAATGCTGCGAGTGTAAATGCGTAACTGTGCTCTCA 1716
 DB 911 ValGlnAsnSerGlyAlaAlaLeuGlnThrSerThrAlaAsnGlyIleGluIleSerGly 930
 QY 1717 CAA-----GCCATTTCTGATCGGATCGGAGGGAATAAATACGACATTAAGAGCT 1764
 DB 931 ThrHisGlnTyrGlyThrPheSerIle---SerGlyAsnLeuAlaThrAsnMetLeuLeu 949
 QY 1765 AAGAGGGGATCAGATTCTCTTT-----AATGATCCCATCGAGATG 1806
 DB 950 GluAsnGlyGlyAsnLeuLeuValLeuAlaGlyThrGluAlaArgAspSerThrValGly 969
 QY 1807 GCAACCGAAATAACAGCCAGCGAGCTTCTCAAACTTCTCAAAATTAACGATGGTGA 1866
 DB 970 LysGlyGlyAlaMetGlnAsnGlnGlyGlnAspSerAlaThrLysValAsnSerGlyGly 989
 QY 1867 GGTATACACA-----GGGATATT 1884

Db 990 GlnTyrThrLeuGlyArgSerLysAspGluPheGlnAlaLeuAlaArgAlaGluAspLeu 1009
Qy 1885 GTTTTGTCTTAATGGAAGCAGTACTTTGTAC----- 1914
Db 1010 GlnValAlaGlyGlyThrAlaIleValTyrAlaGlyThrLeuAlaAspAlaSerValSer 1029
Qy 1915 -----CAAATGTTACATAGAG----- 1932
Db 1030 GlyAlaThrGlySerLeuSerLeuMetThrProArgAspAsnValThrProValLysLeu 1049
Qy 1933 CAAGGAGGATTGTTCTTCGTAAGGCAAAATATCAGTG---AATTCCTAAGTCAG 1989
Db 1050 GluGlyAlaIleArgIleThrAspSerAlaThrLeuThrIleGlyAsnGlyValAspThr 1069
Qy 1990 ACAGGTGGAGTCTGTATATGTAAGCT---GGGAGTACATGGATTTTGTGTAATCCACAA 2046
Db 1070 ThrLeuAlaAspLeuThrAlaAlaSerArgGlySerValTyr----- 1083
Qy 2047 CCACCACACAGCCTCTCGCGCTTAATCAGTTGATCAGCTTTCACATCTGCATTTGTCT 2106
Db 1084 -----LeuAsnSerAsnAsnSerCysAlaGlyThrSerAsnCysGluTyrArg 1099
Qy 2107 CTTTCCTTCTTGTAGCAAC-----AATGCAGTTACGAATCCTCCTACCAATCCTCCA 2160
Db 1100 ValAsnSerLeuLeuLeuAsnAspGlyAsnValTyrLeuSerAlaGlnThrAlaAlaPro 1119
Qy 2161 GCGCAAGATTCTCATCTCGCATCTGTCAGACACTGCTGCTGTTCTGTTCAATATTAGT 2220
Db 1120 AlaThrThrAsn-----GlyIleTyrAsnThrLeuThrThrAsnGluLeuSerGlySer 1137
Qy 2221 GGGCCTATCTTTTGGAGATTGGATGATCAGCTTATGATAGTATGATGATGGCTAGGT 2280
Db 1138 GlyAsnPheTyrLeuHis-----ThrAsnValAlaGlySerArgGlyAspGlnLeuVal 1155
Qy 2281 TCTATACAA-----AAATCAATGCTCCTGAAATTCAGATTAGGAGCTAAGCCC 2328
Db 1156 ValAsnAsnAsnAlaThrGlyAsnPheIlePheValGlnAspThrGlyValSerPro 1175
Qy 2329 CCAGCTAAT-----GAGATGCTAAGTATGGCTAT----- 2385
Db 1176 GlnSerAspAlaMetThrLeuValLysThrGlyGlyAspAlaSerPheSerLeu 1195
Qy 2359 GCGAAT-----CAGATGCTAAGTATGGCTAT----- 2385
Db 1196 GlyAsnThrGlyGlyPheValAspLeuGlyThrTyrGluTyrValLeuLysSerAspGly 1215
Qy 2386 CAAGGAAGCTGAAGCTTCGTTGGAT-----CCTAAT---ACAGCAATAATGGTCT 2436
Db 1216 AsnSerAsnTrpAsnLeuThrAsnAspValLysProAsnProAspProAsnProAsnPro 1235
Qy 2437 TATACTCTGAAGCTACATGGACTAAACTGGGTATATCTCTGGCGCTGAG----- 2487
Db 1236 AsnProAsnProLysProAspProLysProAspProLysProAspProLysProAspPro 1255
Qy 2487 ----- 2487
Db 1256 ThrProGluProThrProThrProValProGluLysArgIleThrProSerThrAlaAla 1275
Qy 2488 -----CGAGTACGCTCTTGTGTTTCA----- 2508
Db 1276 ValLeuAsnMetAlaAlaThrLeuProLeuValPheAspAlaGluLeuAsnSerIleArg 1295
Qy 2509 -----AATAGTTTATGGGATCCATTTTA 2532
Db 1296 GluArgLeuAsnIleMetLysAlaSerProHisAsnAsnAsnValTrpGlyAlaThrTyr 1315
Qy 2533 GATATACGATCT-----GCGCATTCAGCAATTCACCAAGTGTGGATGGG 2577
Db 1316 AsnThrArgAsnAsnValThrThrAspAlaGlyAlaGlyPheGluGlnThrLeuThrGly 1335
Qy 2578 -----CGCTCTATTGTCGAGGATTATGGGTTTCTGAGATT 2613
Db 1336 MetThrValGlyIleAspSerProAsnAspIleProGluGlyIleAlaThrLeuGlyAla 1355

Qy 2614 TCGAATTTCTTCTATCATGACCGCGATGCTTTAGTTCAGGATATCGTATATTAGTGGG 2673
Db 1356 PheMetGlyTyrSerHisSerHisIleGlyPheAspArgGlyHisGlySerValGly 1375
Qy 2674 GGTATTCTCTAGGAGCA---AACTCCTACTTTGGATCATCGATGTTTGGTCTAGCATTT 2730
Db 1376 SerTyrSerLeuGlyGlyTyrAlaSerTrpIleGluSerGlyPheTyrLeuAspGly 1395
Qy 2731 ACCGAAGTATTGCTGATCTAAAGATTATGATGTGTCTTCCATCATCATGCTTTCG 2790
Db 1396 ValValLysLeuAsnArgPheGluSerAsnValAlaGlyLys----- 1409
Qy 2791 ATAGATCCGTTTATCTATCTACCCCAACAGCTTTATGTGTGATCTTATTTGTCGGAGAT 2850
Db 1410 -----MetSerSerGlyGlyAlaAlaAsnGlySerTyr----- 1420
Qy 2851 GCGTTTATCCGTGCTAGCTACGCGGTTTGGGAATCAGCATATGAAACCTCATATACATT 2910
Db 1421 -----HisSerAsnGlyLeuGlyGly---HisIleGluThrGlyMetArgPhe 1435
Qy 2911 GCAGAGGAGCGATGTTGTTGGAT 2937
Db 1436 -----ThrAspGlyAsnTrpAsn 1441
RESULT 28
US-11-052-554A-171
; Sequence 171, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 171
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-11-052-554A-171
Alignment Scores:
Pred. No.: 5,54e-10 Length: 2340
Score: 216.50 Matches: 276
Percent Similarity: 31.9% Conservative: 191
Best Local Similarity: 18.9% Mismatches: 569
Query Match: 2.7% Indels: 426
DB: Gaps: 63
US-10-701-844-1 (1-4435) x US-11-052-554A-171 (1-2340)
Qy 496 GGGGAGAGCTTAACGTGATCATTT---CCCTACTCTGTATAGGATCCGAGTGGGACT 552
Db 105 GlyGluAspLeuAsnThrAsnPheGlyProLeuLysPheIleSerAsnAsnValThrSer 124
Qy 553 ACTGTTTCTCTGCGAGGAGGTTAACTATAAAAAATCTTGAC-----AATTCATTGCA 606
Db 125 IleIleThrGlyValGlyThrLysThrPheSerAsnIleAspPheAlaGlyLysAsnAla 144
Qy 607 GCTTTGCCCTTAAGTGTGTTTGGGAACCTTATAGGAGTTT-----ACTGTTTAGGG 660
Db 145 ThrLeuGlnIleAsnLysAspLeuAsnIleThrThrLysIleAspAsnThrValAlaGly 164
Qy 661 AGAGGACACTCGTTCACCTTTCGAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGT 720
Db 165 AsnAsnGlySerIleThrPheGluGlySerGlyIleIleSerAsnHisIleGlyTyrThr 184

```
QY 721 AATACGCTGCTGATGGACTG----- 741
Db 185 AenSerLeuLeu---GlyLeuAenValGlyAenGlyGluAlaLysIleTyrAlaProGlu 203
QY 742 -----TTTACTATTGAGGTTTAAAGAAATTATCTCTTCCAAATTCGAATTCATTA 792
Db 204 AlaAenAenIleThrIleAenAla---LysAenIleAenLeuThrHisAenAenSerIle 222
QY 793 CTTGCGCTACTGCTGCTCAACGACTAATAAGGGTAGCCAGACTCCGACGCAACATCT 852
Db 223 LeuThrLeuCysAenGlyAenIleThrThrLeuLysGlyAenIleAenAenThrThrGlu 242
QY 853 ACACGCTTAATAGGTACTATTATTTCTTAAACAGATCTTTTGTACTCAATATGAGAAG 912
Db 243 IleAenGlyGlnGlyIleLeuAenLeuAlaTyrAenLeu----- 255
QY 913 TTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTA 972
Db 256 -----GlySerSerSerIleIleThrGlyAenIleGlyAenIleGlySerLeuAen--- 272
QY 973 ACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAAAATACTGCTCAAGCTGATGG 1032
Db 273 ThrIleAenValLeuLeuGlySerAlaThrPheAenSerThrIleLeuLysAlaThrAen 292
QY 1033 GGAGCTTGTAAGTAGTACCAGTTCTCTGCTATGGCTAAACGAGGCTCTATTGCTTT 1092
Db 293 -----IleAenLeuLysHisAenThrSerThrLeuAenLeuAenAenIleIleVal 310
QY 1093 GTAGCGAATGTTGCAGGA-----GTAAGAGGGGGA 1122
Db 311 IleGlyAenIleLysGlyAenAenAenLysAenPheIleLeuAenPheLysValHisGlyThr 330
QY 1123 GGGATTGCTGCTGTTCCAGAGTGGCAGCAGGAGTCTCATCTACTTCAACAGAGAT 1182
Db 331 AenLeu-----AspAenGluMetIleIleProAlaProGlnLysThrHisGly 346
QY 1183 CAGTAGTAGTATGTTTCCAGAAATACTGCGGTAGAGTTGATGGGAGCAGTACCCGAGTA 1242
Db 347 ---ThrLeuAenPheLysGlyAenAlaThrLeu-----AenGlyAenIleAenAenLeu 363
QY 1243 -----CGAGCA----- 1248
Db 364 AenIleLeuLysPheSerGlyGlyHisGlyLysThrLeuAenLeuGlnGlyAenThrLys 383
QY 1249 -----GGG 1251
Db 384 ValAenAenLeuValPheAlaAenSerValLeuAenSerGlyThrIleSerValAenGly 403
QY 1252 ATTTACTCTACGGGAACGTTGCTTTCCTGAAT-----AATGGAAAAACCTGTTT 1302
Db 404 LeuLeuAenThrAenCysValThrPheAenAenSerAenValAenGlyGlyThrLeuIle 423
QY 1303 CTC-----AACATGTTGCTCTCTCTGTTTACATTCGCTCAAGCAACCAACAGT 1353
Db 424 IleAenAlaLysAenThrIleSerAlaLysLeuLeuAenAlaThrLysAlaLysIleGln 443
QY 1354 GGCAGGCTTCTAATACGAGTAATAATTACGGAGATGGAGGA----- 1395
Db 444 IleAenAlaAenLeuThrMetAenHisProSerAlaGlyAenPheSerAenPheIleArgIle 463
QY 1396 -----GCTATCTCTGTAAGAAATGGTCCGCAAGCAGAGTCCATAAC 1437
Db 464 AlaAenAenThrIleTyrThrIleAenPheAenGlyAenGlyAenValAenLeuAenAen 483
QY 1438 TCTGGATCACTTCTCTTGTATGAGAGAGGA-----GTAGTTTCTTTAGTACATGTA 1491
Db 484 AenAlaLysIleIlePheGluGlyAlaAenSerMetLeuAlaLeuIleAenThrGlyVal 503
QY 1492 GCTGCTGGGAAGGGGAGCTATTATTATGCCAAA-----AAGCTC 1530
Db 504 ThrAlaAenArgThrPheThrIleTyrAenAenLeuAenGlnSerGlyAenAenGluTyr 523
```

```
QY 1531 TCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGTAAATATGTTGGTGGAGG 1590
Db 524 GlyIleValLysIleGluAlaIleLysLysValIleThrIleAlaAenGlnSerGlyPro 543
QY 1591 ATTTATTAGAGAAATCTGGAGAGCTCAGTTTATCT-----GCTGATTATGGA 1638
Db 544 TyrThrIleGlyGlnAenAenThrHisArgLeuLysGluLeuIleValGluGlyAlaGly 563
QY 1639 GATATTATTTTCGATGGGAATCTTTAAAGAACAGCAAGAGAGATGCTGCCGATGTTAAT 1698
Db 564 AspIleIleAenAenP----- 569
QY 1699 GCGGTAACTGTCTCTCAAAAGCCATTTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758
Db 570 -----ThrIlePheThrLysLeuLeuSerIleAenSerThrGlyGlnIleThr----- 585
QY 1759 AGAGTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCATCGATGCGCAACCGAAT 1818
Db 586 -----PheAenArgThrLeuAenPheLeuGlyAlaGlyGly 596
QY 1819 AACCAAGCCAGCGAGCTCTTCCAAACTCTTAAAAATTAACGATGGTGAAGGATACACAGGG 1878
Db 597 AenIleAlaPheGlyLysHisGlyThrLeuValValAen-----GlyValThrGly 613
QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGAGCAAGGA 1938
Db 614 SerIleThrThrSerGluAenAenGlnGlyIle----- 624
QY 1939 AGGATTGTTCTTCGTGAAAGCAAAATTTATCGTAATCTCTAAGTCAGACAGGTGGG 1998
Db 625 -----LeuThrIleAenSerGlyAenIleThrGlyVal 635
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCAACAG 2058
Db 636 IleGlyThrAenGluLeuGlyLeuLysLeuValAenIleGlyAlaAenP----- 652
QY 2059 CCTCCTGCGCTAACTCAGTTGATCAGCTTTCCTCAATCTGCAATTTGTCTCTTCTCTTTG 2118
Db 653 -----ValThrCysSer-----AlaAenVal 659
QY 2119 TTAGCAAAATGCGATTACGAATCTCTCTACCAATCTCCAGCGCAAGATTTCTCATCT 2178
Db 660 PheAlaSerValAlaLeuThrAenProSerSerValLeuIleLeuAlaAen----- 676
QY 2179 GCAGTCATTGTTAGCACACTGCTGTTCTGTGTACAAATTTAGTGGGCTATCTTTTGTAG 2238
Db 677 -----GlyValThrLeuThrGlyGluValThr----- 686
QY 2239 GATTTGGATGATACAGCTTATGATAGTATGATTGGCTAGGTTCTAATCAAAAAATCAAT 2298
Db 687 -----HisAenAenThrLysGly 692
QY 2299 GTCTGGAATA-----CAGTTAGGAGTAAAGCCCCAGCT 2334
Db 693 ValLeuSerLeuGlyThrGlySerAenIleThrGlyGlnIleGlyThrAenSerAlaAla 712
QY 2335 AATGCCCATCAGATTGATCTAGG-----AATGAGATCGCTAAGTATCGCTATCAAGGA 2391
Db 713 LeuGluLysIleAenIleGlyAlaGlyAlaSerAenIleAenPheAenIleTyrAlaGly 732
QY 2392 AGCTGGAAGCTTCGCTGGGATCTCTAAT-----ACAGCAAAATAATGGTCTTATCTCTG 2445
Db 733 SerThrValLeuThrAenPheGlnThrSerGluLeuThrLeuAenAenAenValValAen 752
QY 2446 AAGACTACATGGAATAAACTGGGTATAATCTCGGCTCGGAGTAGCTTCTTTGGTT 2505
Db 753 SerAenIleIleThrThrAlaGlyAenAenSerGly-----LysLeuIle 767
QY 2506 CCAATAGTTTATGGGATCCATTTAGATATACGATCTCGGCAATTCAGCAATTCAGCA 2565
Db 768 PheThrGlyAenGlyLysIleThrGlyAenIleGlyAlaAenGlyAlaLeuGlnGlu 787
QY 2566 AGTGTGATGGGCGCTCTTATTCTCGAGGATTTATGGTTTCTCGAGATTTTCGAATTTCTTC 2625
```

```

Db      788 ValVal-----:::|:::|:::|:::|PheAsnGlyThrAsnIle---796
Qy      2626 TATCATGACCGCATGCTTTAGTTCAGGGATATCGGTATATATAGTGGGGGTATATCTTCA 2685
Db      797 -----:::|:::|:::|GlyGlyThrAlaAsn 801
Qy      2686 GGAGCAAACTCTACTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATTTGGT 2745
Db      802 SerGlnAsnPheThrValAlaHisSerAlaAlaAsnValValIleThrGlyLeuThrThr 821
Qy      2746 AGATCTAAAGATTAT-----GTAGTGTCTGCTTCCAACTCATCATCTTGC 2790
Db      822 GlyAlaLeuLysTyrLysAspThrGlyThrIleIleAla-----HisGlyGlyLeu 838
Qy      2791 ATAGGATCCGTTTATCTATCTACCCAAAGCTTTTATGTGATCCTATTTTGTTCGGAGAT 2850
Db      839 ValGlyAspIleAspPheAsnAsnLys-----AlaGlyLysPheIleLeuGlyAsp 855
Qy      2851 GGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTT 2910
Db      856 GlyAlaMetIleAspGlySerValLeuCyAsnGlyGlyValAlaAlaGlyThrLeuAspPhe 875
Qy      2911 GCAGAGGAGCGATGTTTCGTTGGGATAATAACTCTCTGGCTGGAGAGATTGGACGGGA 2970
Db      876 IleGlyAspGlyAsnVal-----ThrGlnAsnIleGlyAlaAsp 888
Qy      2971 TTACCGATTGTGATTACTCCATCTAAGCTCTATTTGAATGAGTTCGCTCTTTCGTGCAA 3030
Db      889 AsnAlaAsnSerIleSerThrIleAsnIleGlnGlyAspAsnThrLysAsnValThrIle 908
Qy      3031 GCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGCA 3090
Db      909 AlaAsnAspIlePheValAspAsnIleHisPheThrAsnGlyGlyLeuGln-----926
Qy      3091 TTCAGAGCGGACATCTCTAAATCTATCATGTTCTCTGTTGGAGTCAAGTTTGATCGATGT 3150
Db      927 ---LeuGlyGlyAsnLeuThrThrHisAsnIleAspPheGlyAlaAsnGlyGlyThrLeu 945
Qy      3151 TCTAGTACACATCTCTAAATATATAGCTTTATGGCGCTTATATCTGTGATGCTTTATCGC 3210
Db      946 GluPheAsnGlyAsnAsnThrTyrAsn---LeuAsnAlaIleValAlaAsnGlyGlnAsn 964
Qy      3211 ACCATCTCTGTGACTGAGACAACGCTCTCTATCCCATCAAGACATGAGACAACAGATGCC 3270
Db      965 GlyIleLeuAsnAlaPheThrAsnLeuLysAlaSerAspAspThr-----979
Qy      3271 TTTTCATTAGCAAGACATCGAGTGTGGTTAGAGGATCTATGTATGCTTCTCTAAACAGT 3330
Db      980 -----IleGlyThrValLysIleIleAsnIleGlnIleGlyThrProGlnAsnPhe 997
Qy      3331 AATATAGAAGTATATGGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGGT 3390
Db      998 ThrIleGlnValAsnAsnLysAsn-----1005
Qy      3391 TTGAGTGCAGGAAGTAGATCGCGTTTCTAAATAATATGTTAGATAGTTAAGTGTAGGG 3450
Db      1006 -----LeuThr 1007
Qy      3451 ATGCGCTTTTCTTTCAGATCTACATCATTTTGTGTTTGTGTTTCTTATTCG 3510
Db      1008 LeuValSerSerValAsnSer-----SerIleAsn 1017
Qy      3511 TATGATTCGCGAGCTCTCTCAAGTGTTAACGCCCTAATGTAAACCATCTCTTTTAAGGGA 3570
Db      1018 PheGlyAspAlaAsnSerGlnLeuIleLeuSerAlaProValAspGlnThrIleLys---1036
Qy      3571 GACGATGTTTACTTGATGGAGACTCGCGCTTTTGTCAATGTCTATGCGAGGACTGAAGAA 3630
Db      1037 -----PheIleAsnAsnLeuAsnGluThrGlyGly 1046
Qy      3631 GGTTCGATT---ATCTCAGCTAATGGCGCAATTTTAAAGATTACCGGACAAAAC-----3681

```

```

Db      1047 GlyIleIleThrLeuAspSerAsnGlyAsnAsnLeuThrIleSerGlyAsnAsnGlyIle 1066
Qy      3682 -----CATACATTATCATTTACAGATTTCTCAAGGCCAGTCTTCTCAA 3723
Db      1067 LysLeuGlySerLysGlyAsnGluLeuSerSerLeuAsnIleLysGlyLysVal-----1084
Qy      3724 AATTATGCTTTCATTTTCAGCAGGAGACACACTTACTCTGAGAGATTTTTCGAGTCTGATG 3783
Db      1085 -----ThrValThrAsnAspLeuAspIleGlnAsnIleHisGlnLeuAsn 1099
Qy      3784 TTCTCGAAAAATGTTTCTTGGCGAGAAAAGGAATGATCTCCGGGAAACCGTGAGTATT 3843
Db      1100 IleAsnAsnGlyAlaLeuPheAspAspGlnSerLeuThrSerAlaLysIleLysAsnIle 1119
Qy      3844 TCCGGAGCAGCGCAAGTATTTCTGGGATAACTCCGTTGGGGTATTCTCTTTTATCTACT 3903
Db      1120 Asn-----IleGlyThr 1123
Qy      3904 GTGCCAACCTCATCACTCCGCTCTCTCCACAGTTAGTGAT-----GCT 3951
Db      1124 Val---AlaGlyGlyAlaIleThrTyrThrLeuAspAlaIleAsnAspAsnPheAspLeuAsn 1142
Qy      3952 CGGAAGGGTCTATTTTCTGTAGAGACTAGT---TTGGAGATCTCAGGGGTCAAAAAA 4008
Db      1143 ThrSerGlyMetValPheLysHisGlnAspSerIleLeuGluLeuLysAsn-SerSerAs 1162
Qy      4009 GGGGTCA-----TGTTTCATTAATATGCGGGGAATTTTCGGAACAGTT 4050
Db      1162 nThrAsnAspHisThrIleThrLeuThrSerAlaLeuAspProGly---AsnAsnGlnPhe 1181
Qy      4051 TTTCGAGTAAGAATAATAATAATGCTGCTGGTGGAGGACGTGGGTTCCGCTACACCATC 4110
Db      1181 eGlyIleIleLysLeuIle-----ThrAs 1189
Qy      4111 AAGTACGACTTTTACAGTTAAAAAAGTGTAAAGGAAAGTTTCTTTCACA-----4159
Db      1189 pThrAsnLysLeuThrIleAspAsn---AsnGlyAsnValAlaTyrThrLeuGlyThrAl 1208
Qy      4160 -----GATTAACGTAGCC---TCTTGGCGG 4179
Db      1208 aAsnHisMetLeuLysGlnLeuThrPheAlaSerIleAspAsnGlyAlaIleAlaLeuLys 1228
Qy      4180 AGGCGAGTGTGTTTATAAGGCATTTGTGCTT---TTCAAAGACAATGAA-----4225
Db      1228 sValGlyIleAsnValGluAsnValThrLeuAsnIleLysAspIleGluLeuAsnGluVal 1248
Qy      4226 ---GGAGGCATATTTCTCCGAGGGAACACAGCATACGATGATTTTAAGGATTTCTGTGTC 4281
Db      1248 LAsnAlaAsnValLeuPheAsnLysAsnThrThrTyr-----ThrAlaTh 1263
Qy      4282 TACTAATCAGGATCAGAATACGAGACAGGAGGCGGTGAGAGTATTATTGCTCTCCAGA 4341
Db      1263 rGlyAsnIleAsnGlyHisValAspPheGlnGlyAsnAlaGlyValIleAsnLeuAsnAs 1283
Qy      4342 TGAT-----TCTGTAAAAGTTTGAAGGCAAT---AAAGGTTCTATTGTTT 4383
Db      1283 pAspIleGluLeuAspGlySerValThrSerThrGlyAsnValAsnGlyThrLeuAsnPh 1303
Qy      4384 TGAT 4387
Db      1303 eAsn 1304

```

RESULT 29

```

US-11-052-554A-283
; Sequence 283, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; CURRENT FILING DATE: 2005-02-07

```

;; PRIOR APPLICATION NUMBER: US 60/589,227
;; PRIOR FILING DATE: 2004-07-20
;; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;; PRIOR FILING DATE: 2004-02-06
;; NUMBER OF SEQ ID NOS: 763
;; SOFTWARE: Patent in version 3.3
;; SEQ ID NO 283
;; LENGTH: 1461
;; TYPE: PRT
;; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-283

Alignment Scores:
Pred. No.: 9,99e-10 Length: 1461
Score: 212.50 Matches: 295
Percent Similarity: 33.7% Conservative: 200
Best Local Similarity: 20.1% Mismatches: 540
Query Match: 2.7% Indels: 432
DB: 11 Gaps: 75

US-10-701-844-1 (1-4435) x US-11-052-554A-283 (1-1461)

```
QY 439 TCTTTAAATGGGGGGATATGCA---GCAGAAATCATGGTTCCTCAAGGAATTTACGAT 495
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 ThrPheSerGlySerGlyGlnAlaGlyAlaThrIleGlnIleLysAspSer-----Asn 220
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GGGGAGACGTGA-----ACTGTATCATTTCCC 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 GlyAsnThrIleAlaSerThrGlnValAspAsnAsnGlyHisTrpSerValSerLeuPro 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 TATACGTATTAGGAGATCCGAGT-----GGGACTACTGTGTTTT 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ThrGlnSerAlaGlyGluHisThrTrpSerValGlnIleValGlySerThrIleThr 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 TCTGAGAGAGTTAACTAAATAAACTTGACAAATCTTGACAAATCTTGAGCTTTGAGCTTAACT 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 AspAlaGlySerIleThrLeu---ThrIleAspAsnSerGlnAlaSerValGlnValAla 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 622 TGTTTTGGG-----AACTATTATTA 639
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 ThrThrAlaGlyAspAsnIleIleAsnAlaSerGluGlnAlaAlaGlyPheThrLeuSer 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 GGGAGTTTACTGTTTAGGGAGAGACACTCGTTGACCTTCGAGAACATACGAGCTTCT 699
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 GlyThrSerSerHisLeuAlaGlnGlyThrGluLeuThr-----ValThr 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 ACAATATGGGCGAGCTCTAAGTAATAGC---GCTGCTGATGAGCTGTTTACTATTGAGGCT 756
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 LeuAsnGlyLysThrTyThrThrSerValGlyAlaAsnGlyAlaTrpSerValGln--- 333
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 TTTAAGAATTTATCCTTTTCCAAATTCGAATTCATTACTTGGCGTACTGCTGTCMACG 816
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 -----ValProThrAlaAspAlaGlnAla 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 ACTAATAAGGTAGCCAGACTCCGACGACAACTACACCGCTCTAATGCTACTATTAT 876
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 LeuGlyGluGlyAsnGlnAlaValLeuValSerGlyLysAspAlaThrGlyAsnThrVal 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 TCTAAAACAGATCTTTGTTACTCAATAATGAGAAGTCTCTCAATCTTATGTAATTTAGTC 936
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 ThrGlyAlaGlnLeuLeuThrValAspThrGlnProProThrLeuAlaIleAsnThrIle 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 TCTGGAGAT-----GGGGAGCTATAGATGCTTAAGAGCTTAACGGTTACAGGAAT 987
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 AlaGlnAspAsnIleIleSerAlaAlaGluHisAsnValAlaLeuValLeuSerGlyThr 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 AGCAAGCTTTGTCTCTTCCAGAAATATCTGCTCAAGCTGATGGG-----1032
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 SerAspAlaGluAlaGlyGlnThrValThrLeuThrValAsnGlyLysSerHisThrAla 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1033 -----GGAGCTTGTCAAGTA-----GTCACCAAGTTTCTCTGCTATG 1068
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 ThrValGlySerAspGlyThrTrpGlnValThrLeuProAlaThrGluValGlnAlaLeu 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 1069 GCT-----AACGAGGCTCTATTGCTTTCAGCAATGTTGACGAGTAAAGGGGGA 1122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 AlaGluGlyAsnTyAlaValAsnAlaSerValSerAspArgAlaGly-----457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1123 GGGATTGCTGTTGAGGATGGCAGCAGGAGTGTCTATCTACTTACACAGAAGAT 1182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 -----AsnThrThrSerHisSerAlaAsnPheThrValAspThrSerAla 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1183 CCAGTAGTAAGTTTTTCCAGAAATATCTGCGGTAGAGTTTGTAGGAACTAGCCCGAGTA 1242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 ProValValSerVal-----AsnThrValAlaGlyAspAsp-----484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1243 GGAGGAGGATTTACTCTACGGGAACGTTCTTCTTGAATAATATGAAAAACCTTGT 1302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 -----IleLeuAsnAsnAlaGluGln-----491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1303 CTCACAATGTTGCTTCTCTGTTTACATTCTGTAGCAACCAACCAAGTGGCAGGCT 1362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 -----AlaValAlaGlnIleIleSerGlyGlnVal 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1363 TCTAATACGAGTAATAATTACGAGATGGAGGAGCTATCTTCTGTAAGAATCGTCGCAA 1422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 SerGlyAlaSer-----ProGlyAspThrValThrValLysLeuGlyThrHis 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1423 -----GCAGGATCCAATAACTCTGGATCAGTTTCTCTTGTAT-----1458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 518 ValLeuThrGlyIleValAlaLeuAspGlySerTrpAsnValAlaLeuAspProAlaVal 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1459 -----GGAGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTCTGGG 1500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 ThrArgThrLeuAspArgGlyAlaAsnThrIlePheValThrValThrAspAlaAlaGly 557
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1501 AAAGGGGAGCTATTATATGCCAAAAGCTCTCGGTTGCTAACTGTGGCCTGTACAATT 1560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 AsnThrGlyAlaAla---SerArgAlaIleThrLeuValGlyValSerProLeuIleThr 576
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1561 TTAAGGAATATCGCTTAATGATGGT-----1584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 IleAsnThrValSerGlyAspAspIleIleSerGlyAlaGluLysGlyAlaProLeuThr 596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1585 -----GGAGCGATTTATTAGGAGAACTCGAGAG-----CTCAGTTTA 1623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 597 LeuThrGlySerThrGlnGlnAlaGluThrGlyGlnThrValThrValThrLeuAlaGly 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1624 TCTGCTGATTATGGAGATATTATTTTCATGGGAATCTTAAAGAACAA-----1671
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 GlnSerPheThrThrValGlnAlaAspGlySerTrpSerLeuThrValProAlaAla 636
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1672 GCCAAAGAGAAATGCTGCCGAT---GTTAATGGCGTAACCTGTCTCTCAACGCCATTCG 1728
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 AlaMetGlyAsnLeuProAspGlyAlaValAlaIleThrAlaSerValThrAspLeuSer 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1729 ATGGATCGGGAGGAAAAATACGATTAAGAGCTTAAGCAGGAGGCGCATCTCTCTTT 1788
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 GlyAsnThrGlyAsnThrSerArgThrIleThrValAspSerGlnAlaProAlaLeuSer 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1789 AATGATCCCATCGATGGCAAC-----GGAAATACCCAGCCA 1827
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 IleAspProLeuThrAlaAspAsnIleIleAsnAlaAlaGluSerGlyGlnAspLeuPro 696
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1828 -----GCGCAGCTTCTCCAAACTTCTAAAAAATTAAC---GATGGTGAA 1866
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 IleThrGlyThrThrAspAlaGlnProGlyGlnThrValThrValThrLeuAsnGlyGln 716
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1867 GGATACACACAGGGGATATT-----1884
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 ThrTyThrGlnGlyValValGlnProAspGlyThrTrpSerValThrValProAlaAlaAsn 736
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1885 -----GTTTGTCTTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCA 1938
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 ValGlyAlaLeuAlaAspGlyAsnAlaThrValThrAlaSerValAsnAspValAlaGly 756
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

1939 QY -----AGGATTGTTCTT----- 1950
757 Db AsnProSerSerValSerArgValAlaLeuValAlaThrProProValValThrIle 776
1951 QY -----CGTGAAGGCGAAATATCATCGTGAAT 1977
777 Db AsnProValAlaThrAspAsnValIleAsnThrProGluHisAlaGlnAlaGlnIle 796
1978 QY TCCTTAAGTCAGACAGCTGGAGCTGTATATGGAAGCTGGAGTACATGCGATTTTGTA 2037
797 Db SerGlyThrValThrGlyAla-----GlnAlaGly-----AspIleVal 809
2038 QY ACTCCACCAACCAACACAGCCTCCTGCCGCTAATCAGTTGATCAGCGCTTCCAAATCTG 2037
810 Db Thr-----ValThrLeuAsnAsnVal 816
2098 QY CATTTGCTCTTTCTTCTTTGTTAGCAACAATGAGTGTAGCATCTCTCCACCAATCCT 2157
817 Db AspTyrThrThrValValAspGlySerGlyAsnTrpSerLeuGlyValProAlaSerVal 836
2158 QY CCA-----GGCAAGATTCTCATCTCGAGTCATT-----GGT 2190
837 Db ValSerGlyLeuAlaAspGlySerTyrProValSerValSerValThrAspIleAlaGly 856
2191 QY AGCAACTGCTGCTGTTCTGTACAAAT-----AGTGGCGCTATCTTTTGGAG 2238
857 Db AsnThrGlySerGlnSerLeuThrValThrValAsnThrAlaAlaProLeuIle----- 874
2239 QY GATTGGATGATACAGCTTATGATAGGTATGATGGTGTAGTGTCTTAATCAAAAAATCAAT 2298
875 Db GlyIleAsnSerIleAlaGlyAsp-----AspValIleAsnAlaSerGluLysGlyAla 892
2299 QY GTCCTGAATATCAGTATAGGCAAGCCCGCAGCTAATGCCCATCAGATTGACTCTA 2358
893 Db AspLeuGlnIleThrGlyThrSerAspGlnProValAsnThrAlaIleThrValThrLeu 912
2359 QY GGG-----AATGAGATGCTAAGTATGGCTATCAAGGAAGCTGGAAGCTT----- 2403
913 Db AsnGlyGlnAsnTyrThrThrThrAspAlaSerGlyAsnTrpSerValThrValPro 932
2404 QY CGCTGGGATCCTTAATACAGCAAAATATGGTCTTATCTCTGAAGCTACATGGACTAAA 2463
933 Db AlaSerAlaValThrAlaLeuGlyGlnAlaAsnTyrThrValThrAlaAlaValThrSer 952
2464 QY ACTGGGTATATCTCGGCGCTCAGCGAGTGTCTTGTGTTCCAAATGTTTATGGGA 2523
953 Db AspIleGlyAsnSerAlaThrAlaSerHisAsnValLeuValAspSerAlaLeuProGly 972
2524 QY TCCATTTTAGAT-----ATACGATCTGCGCATTCAGCAATTCAA 2562
973 Db ValThrIleAsnProValAlaThrAspAspIleIleAsnAlaAlaGluAlaGlyValAla 992
2563 QY GCAAGTGTGATGGCGCTCTTATTGTGAGGATATGGGTTCTGGAGTTTCGAATTC 2622
993 Db GlnThrIleSerGlyGln-----ValThrGlyAla----- 1002
2623 QY TTCTATCATGACCGGATGCTTTAGTGTAGGATATCGGTATATAGTGGGGTTATCC 2682
1003 Db -----GluAspGlyAspThrVal-----ThrIleThr 1011
2683 QY TTAGGAGCAACTCTCTACTTTGGATCATCGATGTTTGTGTAGCATTTACCGAAGTATT 2742
1012 Db LeuGlyGlyAsnThrTyr-----ThrAlaThrVal 1021
2743 QY GGTAGATCTAAAGATTATGATAGTGTGCTGTTCAATCATCATGCTTGTGCATAGATCCGTT 2802
1022 Db GlySerAsnLeuThrTrpSerVal-----AspVal 1031
2803 QY TATCTATCTACCCCAAGCTTTATGTGATCTCTATTGTTGCGAGATGGTTTATCCGT 2862
1032 Db ProAlaAlaAspIleGlnAlaLeu-----GlyAsn----- 1047
2863 QY GCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATATATTTGCGAGGAGAGC 2922

1048 Db AlaSer-----ValThrAsnGlnAsnGlyAsnThrGly-----SerGlyThrArg 1062
2923 QY GATGTTCTGTTGGATATAAAC-----TGTCGTGCTGGAGAG--- 2958
1063 Db AspIleThrIleAspAlaAsnLeuProGlyLeuArgValAspThrValAlaGlyAspAsp 1082
2959 QY -----ATTGAGCGGATATACCATGTGTGATTACTCCATCTAAGCTCTATTG 3006
1083 Db ValValAsnIleIleGluHisGlyGlnAlaLeuValValThrGlySerSerGlyLeu 1102
3007 QY AATGAGTTGGTCTCTTC-----GTGCAA 3030
1103 Db AlaGluSerThrProLeuThrValThrIleAsnAsnValGluTyrThrThrAlaValGln 1122
3031 QY GCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGGCA 3090
1123 Db AlaAspGlySerTrp-----SerValGlyValThrAlaAlaGlnValSerAla 1138
3091 QY TTCAAGAGCGGACATCTCTAAATCTATCAGTT-----CCTGTTGGAGTGAG 3138
1139 Db TrpProAlaGly---ThrValAsnIleAlaValSerGlyGluSerSerAlaGly----- 1155
3139 QY TTTGATCGATGTTCTTACATACATCTTAATAATATAGCTTTATGGCGCTTATATGCT 3198
1156 Db ---AsnSerValSerIleThrHisProValThrValAspLeuThrProAlaAlaIle--- 1173
3199 QY GATGCTTATCGCACCATCTCTGGTACTGAGACAACGCTCTATCCCATCAAGAGACATGG 3258
1174 Db ---ThrIleAsnThrIleAla----- 1179
3259 QY ACAACAGATGCTTTTCATTTTACGAAGACATGAGATTGTGTTAGAGGATCTATGATGCT 3318
1180 Db ThrAspAspValIleAsnAlaAlaGluLysGlyAlaAspLeu-----ThrLeuSer 1196
3319 QY TCTCTAACAGTATATAGAA-----GTATATGGCCATCGAAGA 3357
1197 Db GlyThrThrAsnValGluProGlyGlnThrValThrValThrPheGlyGlyLeu 1216
3358 QY TATGAGTATCGAGATGCTCTCTCGAGGCTATGTTTGTAGTCGAGGAAGTAGAGTCGGTTC 3417
1217 Db TyrThrAlaSerValAlaSerAspGly----- 1225
3418 QY TAAATAATTTGTTAGATAGTTAAGTGTACGATGCTTTTCTTTGAGATCTACATCA 3477
1236 Db ---SerTrpThrAlaThrValProAlaAlaAspLeuAlaSerLeuProGluGlySer 1243
3478 QY TTTTGTGTTTGTAGCTGTTTGTGTTCTTATTCGTATGATTTCGCGAGCTCTCTCAAGTG 3537
1244 Db AlaSerAlaLeuAlaSerValSerAsnIleAsnGlyAsnSerAlaSerAla----- 1260
3538 QY TTAACGCTAATGTAAACCACTCTTTTAAAGGAGACGATGTTTACTTGAATGGAGACTGC 3597
1261 Db ---ValHisAsnTyrSerValAspSerSerAlaProThrIleIleIleAsn---ThrVal 1278
3598 QY GCTTTGTTCAATGCTATCTCAGGAGCTGAAGAAGTTTCGATTCATCTACGCTAATGGCAG 3657
1279 Db AlaSerAspAsnIleValAsnAlaSerGluAlaAspAlaGlyValThrValSerGlySer 1298
3658 QY AATTAAAGATTACCGGACAAACCATATCATTTATCTTACAGATCTCAAGGGCCAGTT 3717
1299 Db ThrThrAlaGluAlaGlyGlnIleValThrIleThrLeu-----AsnSerProThr 1315
3718 QY CTTCAAAATAT---GCCTTCATTTTACGAGGAGAGACACTTACTCTGAGAGATTTTTCG 3774
1316 Db ValGlnThrTyrGlnAlaThrValGlnAlaAspGlySer----- 1328
3775 QY AGTCTGATGTTCTCGAAAAATGTTTCTTCGCGAGAAAAGGGAATGATCTCCGGGAAAACC 3834
1329 Db ---TipSerIleAsnIleProAlaAlaAspLeuGluAlaLeuThr----- 1342
3835 QY GTGAGTATTTCCGGAGCGCGCAAGTGATTTTCTGGATAACTCTCGTGGGGTATTCTCTCT 3894

```

Db 1343 -----AspGlySerHisThrLeuThrAlaThrValAsnAspLysAlaGlyAsnPro 1359
QY 3895 TTATCTACTGTGCGCAACCTCATCA-----TCACCTCCCGCTGCTCCACAGATTAGTGAT 3948
Db 1360 AlAserThrThrHisAsnLeuAlaValAspLeuThrValProValLeuThrThrIleAsnThr 1379
QY 3949 GCTCGGAAGGGTCTATTTTCTGTAGAG-----ACTAGTTTGGAGATCTCAGGC 3999
Db 1380 IleAlaGlyAspAspIleIleAsnAlaThrGluHisGlyGlnAlaLeuValIleSerGly 1399
QY 4000 GTCAAAAAAGGG-----GTCATGCTTCGATAATAATCCCGGAATTC 4041
Db 1400 SerSerThrGlyGlyGluAlaGlyAspValValThrValThrLeuAsnSerLysThrTyr 1419
QY 4042 GGAACAGTTTTCGAGGTAGGAATAATAATATGCTGGTGGG-----GA 4086
Db 1420 ThrThrThrLeuAspAlaSerGlyAsnTrpSerValGly-ValProAlaAlaAspValTh 1439
QY 4087 GGCAGTGGGTCCGCTACACCACTCAAGTAGCACTTTTACAGTTTAAAAAACTGTAAAGGGAA 4146
Db 1439 rAlaLeuGlySerGly-----ProGlnThrValThrAlaThrValThrAspAlaAlaGlyAs 1458
QY 4147 AGTTTCTTTCACAGATAAC 4165
Db 1458 n-----SerAspAsn 1461

RESULT 30
US-11-087-099-7521
; Sequence 7521, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7521
; LENGTH: 7465
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-7521

Alignment Scores:
Pred. No.: 5.19e-09 Length: 7465
Score: 208.00 Matches: 313
Percent Similarity: 29.4% Conservative: 169
Best Local Similarity: 19.1% Mismatches: 617
Query Match: 2.6% Indels: 541
DB: 11 Gaps: 68

US-10-701-844-1 (1-4435) x US-11-087-099-7521 (1-7465)
QY 451 GGGGATATGACAGCAAGATC-----ATGGTTCTCAAGAAATTTACGATGG 498
Db 5670 GlyAlaAspThrAlaGluValValSerIleAsnLeuThrAspGlnThrLeuPheAlaGly 5689
QY 499 GAGACGTTAACTGTATCATTTCCCTACTACTGTTATAGGAGATCCGAGTGGGACTACTGTT 558
Db 5690 AspSerValAspIleThrPheSerArgAla----- 5699
QY 559 TTTTCTCGCAGAGATTAAACATTTAAAAATCTTGACAAATCTTATGAGCTTTGCCCTTAA 618
Db 5700 ---LysThrGlyAsnLeuThr-----SerIleThrSerThrGlyLeu 5712
QY 619 AGTTGTTTTCGGAATTTATTAGGAGTTTACTGTTTATAGGAGAGACACACTCGCTGACT 678
Db 5713 TyrPheGlyGlyAsn-----AspAlaLeuValValGlyAspGly----AspValThr 5728
QY 679 TTC-----GAGACATACGGACTTCTCAAAATGGGCA----- 711
Db 5729 PheIleGlyGlyLysAspAsnAspSerLeuIleValAspSerThrGlyAlaThrThrAla 5748

```

```

QY 711 ----- 711
Db 5749 GluValLeuAlaGlyThrAlaGlyAspProProAsnGlnArgPheAlaPheGlyAspAsn 5768
QY 712 -----GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTGAG--- 753
Db 5769 ValTyrLeuGluTyrSerLeuAspAspAlaGlyLysGluAlaLeuIleThrIleAsnSer 5788
QY 754 -----GTTTTAAAGAAATTATCTCTTTTCCCAATTCGAATTTCA 789
Db 5789 GluAspThrThrAlaSerThrGlyAspAspSerLeuLeuIleGlyProThrGlySer 5808
QY 790 TTATCTTGGCGTA-----CTCCGCTGCTCAACGACTAAT 822
Db 5809 LeuAlaGlyValMetGlyGluValIleAlaValGlyGlyMetGlyAlaAspThrIleAsn 5828
QY 823 AAGGTAGCCAGCACTCCGACGACACATCTACCGCTCTAATGGTACTATT--- 873
Db 5829 ValSerGlyLeuThrAlaAspThrValAlaAlaGlyAspAsnIleAsnIleThrArgSer 5848
QY 874 -----TATTCTAAACACAGATCTTTTGTACTC-----AATAAT 906
Db 5849 ThrGluGlyTyrLeuLysThrMetArgSerLeuLeuProValGluGlyGlyValAspAsn 5868
QY 907 GAGAAGTTCTCATTTCTATAGTAATTTAGTC-----TCTCGAGATGGGGAGCT 954
Db 5869 IleSerLeuSerAlaGlyAspHisValValIleGlyGlyMetSerGlyAspThrIleLeu 5888
QY 955 ATAGATGCTAAGAGCTTAACGGTTTCAAGAAATTAGCAAGCTTTGTCTCTCCAGAAAT 1014
Db 5889 ThrGlySerGlyAsnThrIleAlaMetGlyAspSerGlySerLeuValPheAspSerAsp 5908
QY 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAA-----GTAGTCACAGATTC 1059
Db 5909 ThrAlaProValTyrTrpGlySerValAspSerThrAspAspProThrLeuThrGluAla 5928
QY 1060 TCTGCTATGGCT-----AACGAGGCTCTTATTGCTTCTTGTAGCGAATGTGCA 1107
Db 5929 GlnGluGlnAlaValLysArgValAsnArgTrpGlyIleLeuGlnThrLeuSerSerLeu 5948
QY 1108 GGAGTAAGAGGGGAGGATGCTGCTGTTTCAGGATGGGACGAGGAGGTGTCATCATCT 1167
Db 5949 GlyThrThrAsnGlyGlyHisAspThrMetThrLeuGlyGlnLysGlyThrSer----- 5966
QY 1168 ACTTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATATCTCGGTAGAGTTTGTATGG 1227
Db 5967 -----PheAspGly 5969
QY 1228 AACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTGCTTCTTGAATTAAT 1287
Db 5970 AsnLysValAlaValGlyGly----- 5976
QY 1288 GGAACAACTGTTTCTCAACAATGTGCTTCTCTGTTTATCATTTGCTGCTAAGCAACCA 1347
Db 5977 -----IleGlyAsnAspThrIleThrLeuTyrGlyThrArgSerTyrVal 5991
QY 1348 ACAAGTGGACAGCTTCTAATACAGTAATAATTAATTCAGGATCGAGAGCTATCTTCTGT 1407
Db 5992 GluThrAspAlaAlaGlyAsnAlaValGluAsnValAspGlySerAsnThrIleValVal 6011
QY 1408 AAGAATGGTGGCNAAGCAGGATCCAATACTCTCGGATCAGTTTCTTGTAGAGAGGGA 1467
Db 6012 ThrGluArgAlaValAlaGlyAspAsn-----AlaThrMetSerPheGlyAlaAspGly 6029
QY 1468 GTAGTT-----TTCTTTTAGTACATGTAGCTGCTGGGAAGGGGAGGACTATT 1515
Db 6030 LeuIleThrAspPheAlaThrGlyAspSerAsnIleAlaThrAlaGlyAsnAspThrIle 6049
QY 1516 TATGCCAAAAGCTCTCGTTGCTTAAGTGGCCCTGTACAAATTTTAAAGAAATATCGCT 1575
Db 6050 -----LysValAspValSerAsn---AspProValIleAspProAspTyrValAsp 6065
QY 1576 AATGATGGTGGAGCGATT-----TATTTA-----GGAGAA 1605

```


Db	6066	GlyAspIleGlyThrLeuAspMetTyrLeuThrThrGluPheAsnValLeuAlaGlyGly	6085
Qy	1606	TCTGGA-----GAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGAT	1653
Db	6086	LeuGlyAsnAspValLeuGluValValGlySerThrGlyMetArgGluValIleSerGly	6105
Qy	1654	GGGAATCTT-----AAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAGCGCTA	1704
Db	6106	AspAsnLeuLeuTyrGlnArgGlyThrAlaSerGlnAsnTyrAsnHisLeuTyrAlaGlu	6125
Qy	1705	ACTGTGCTCTCACAAGCC-----ATTTGATGGATCGGAGGGA--- 1746	
Db	6126	VallysSerSerSerGlyGlyThrAspThrIleLeuThrGlyAsnGlyGluLeu	6145
Qy	1747	-----ATAACGACATTAAGAGCTAAAGCAGCGGCGATCAG	1779
Db	6146	IlePheGlyGlyValGlyAsnAspThrLeuThrValLeuThrThrAlaValAspHisSer	6165
Qy	1780	ATTCTCTTTAATGATCCCATCGAGATGCGAAACGGAATAACAGCCAGCGAGCTTCC	1839
Db	6166	IleValPheGlyAspAlaGlyIysValAlaTyrAspThrAsnAlaThrGlyAlaValSer	6185
Qy	1840	AAACTCTTAAAAATT-----AACGATGGTGAAGG-----TACACAGG	1878
Db	6186	LeuIleAlaSerIleProGluThrAspAspGlyAsnAspThrValThrVallysAsnGly	6205
Qy	1879	GATATTGTTTGGTAAAGCAGCTACT-----TTGTACCAAAATGTTACGATAGAG	1932
Db	6206	AspLeuTyrLeuPheGlyGlyGlyThrAspAlaLeuThrValSerAlaAsnAspThr	6225
Qy	1933	CAAGGAAGATTGTTCTGTGGAAGCAAAATATCATGTAATTCTCTAAGTCAGACA	1992
Db	6226	AlaThrArgValValMetGlyAspSerGlyGlnValAsnMetThrLeuTyrSerTyrThr	6245
Qy	1993	GGTGGAGTCTGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACACCA	2052
Db	6246	AspAla-----AspGlyValThr----- 6251	
Qy	2053	CAACAGCTCTCGCGCTAAATCAGTTGATCAGCTTCCAAATCTGCATTTGCTCTTCT	2112
Db	6252	-----GlnThrIleGlyValProAsnPhe-----IleGlnValAsn	6263
Qy	2113	TCTTTGTAGCAACAATGCAATCTCCTACCAATCTCCAGCGCAGATCT	2172
Db	6264	AspAlaSerThrThrAsnGluSerThrArgIleAspThrPheThrLeuProSerIysGly	6283
Qy	2173	CATCTCGAGTCATTGGTAGCAACTGCTGCTTCTGTACAAATAGTGGGCTATCTTT	2232
Db	6284	AspAsnLeuIleIleGly-----LeuGlyThrAspLeuLeuSerGly-----	6298
Qy	2233	TTTGAGGATTGGATGATACAGCTTATGATAGTATGATGGTAGTTCTAATCAAAA	2292
Db	6298	----- 6298	
Qy	2293	ATCAATGCTCGAAATTACAGTTAGGACTAAGCCC-----CCAGCTAATCCCATCA	2346
Db	6299	-----AlaGlnLeuGlyValAspSerIleIleProGlyThrGlySerIle	6313
Qy	2347	GATTGACTTAGGATGAGATGCTTAAGTATGC----- 2382	
Db	6314	GlnValSerLeuAlaAsnSerAlaAspProThrGlyAlaTyrSerIysValValSerVal	6333
Qy	2383	-----TATCAAGGAAGCTGGAAGCTTGGCTGG	2409
Db	6334	ValValLeuGlyGlnTyrAsnGluMetGlyIleThrAspGlyThrTyrValLeuSerGln	6353
Qy	2410	GAT-----CCTAATACA-----GCAAAATAAATGCTTATCTAATCTGAAA	2448
Db	6354	AspGlyValAspThrArgProGluThrThrAlaThrGlySerGlyGluValThrGluAsp	6373
Qy	2449	GCTACATGACTAAACCTGGTAT-----AATCTGGGCTGAGCGA	2490
Db	6374	GlyThrGlnSerValThrGlyTyrLeuGlyLeuAspAlaLeuAsnGlyGlyLeuAlaIle	6393
Qy	2491	GTAGCTTCTTTGGTCCAAATAGTTTATGGGATCCATT----- 2529	
Db	6394	PheThrGluAlaThrThrThrLeuLeuTyrGlyThrLeuThrValIysAspGlyGly	6413
Qy	2530	-----TTAGATATACGATCTGCGCATTCAGCAATTCAGCA----- 2565	
Db	6414	TrpThrTyrThrLeuAsnAsnSerSerAlaValGlnAlaLeuMetSerGlyAsnAsp	6433
Qy	2566	-----AGTGTGATGGCGCTCTTATTGTCGAGGATATCG	2601
Db	6434	ArgProGluThrPheGlnValAsnThrThrAspGlyGlnGlnThrThrValThrIleArg	6453
Qy	2602	GTTCCTGGAGTTTCCAAATTTCTTCTATCATGACCGCGATGCTTTAGGTCAGGATATCG	2661
Db	6454	VallysGlyAlaAspAsp-----ValSer	6461
Qy	2662	TATATTAGTGGGTTTATCTCTTAGGACCAACTCTTCTTGGATCATCGATGTTTGT	2721
Db	6462	ThrLeuThrGlySerSerSerAlaSerLeuAsnGluAspThrSerAlaValSerGly	6481
Qy	2722	CTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCCAATCAT	2781
Db	6482	---ThrLeuThrValValAspAlaAspValIleAspAlaThrValThrAlaAlaThrSer	6500
Qy	2782	CATGCTTCATPAGGATCGTTTATCTATCTACCCAA----- 2817	
Db	6501	ValGlyThrTyrGlyThrPheSerValGlySerAsnGlyValTyrSerTyrGlnLeuAsp	6520
Qy	2818	-----CAAGCTTATGTGGATCTTATTTGTCGGAGATGCGTTATC	2859
Db	6521	AsnSerIysAlaValValGlnGlyLeuThrGlnGlyGlnValSerGluSerPheThr	6540
Qy	2860	CGTGTACTACGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAG	2919
Db	6541	ValHisThrSerAspGlyGlySerGlnSerLeuThrMetAsnIleThrGlyArgGlnAsp	6560
Qy	2920	AGCGATGTTCTGGGATTAATCTCTGCTGCGAGAGATTGGAGCGGATTAACCGATT	2979
Db	6561	SerAlaVal-----ValGlySerGlyThrGlySer	6570
Qy	2980	GTGATTACTCTAAGCTCTATTTGAATGAGTTGCTCTTCTGTCAGCTCAGTTT	3039
Db	6571	ValThrGluAspSerThrLeuSerSerGlyMetLeuSerValSerAspAlaAspSer	6590
Qy	3040	TCTTATGCGATCATGAATCTTTACAGAGCAAGC-----GAT	3078
Db	6591	GlyGluAlaAlaTyrValAlaSerThrGlnThrGlySerTyrGlySerPheSerLeuAsp	6610
Qy	3079	CAAGCTCGGCATTCAGAGCGGACATCTCTTAATCTATCAAGTTCTCTGTTGAGTGAAG	3138
Db	6611	SerSerGlyAlaTyr-----SerTyrGlnLeuAsn----- 6620	
Qy	3139	TTTGATCATGTTCTAGTACATCTCTAATAATATAGTCTTATGCGGCTTATATCTGT	3198
Db	6621	-----AsnSerAlaSerGlnValGlnAlaLeuMetGlyGlnGlnValSer	6636
Qy	3199	GATGCTTAT-----CGCACCATCTCTGCTACTGAGACA-----ACGCTCTCTA	3240
Db	6637	GluSerPheThrValAlaThrValAspGlyThrAlaSerSerValValValThrValVal	6656
Qy	3241	TCCCATCAAGACATGAGCAACAGATGCTCTTTTCAATTTAGCA-----AGACATGA	3291
Db	6657	GlyAlaGln-----AspAlaAlaGlnLeuGlyGlyValGlnSerGlyGly	6671
Qy	3292	GTGTGGTTAGGGA-----TCTATGATGCTCTCTTAACCAAGTAATATAGATATAT	3345
Db	6672	ValValGluAspGlyGlnLeuSerThrAlaGlyGlnLeuThrIleSerAspAlaAspAla	6691
Qy	3346	GGCCATGGAAGATGAGTATCGAGATGCTTCCGAGGCTATGTTTTCAGTGCAGGAAGT	3405
Db	6692	GlyGlnSerLeuTyrGlnValAlaSerThrAspGlyMetTyrGlyHisPheAlaMetAsp	6711

QY	3406	AGAGTCCGGTTCTAAATAATTCGTTTAGATAGTTAAGTGTTAGCGATGCCCTTTTCTTGTG	3465
DB	6712	SerSerGlyAlaTrpSerTyValLeuAsnAsnGluLeuAlaIleGlnMetIle	6731
QY	3466	AGATCTACATCATTTGTTTTTTAGCTGTGTTGTGTTCTTATTCTGATGAATCGCGAGC	3525
DB	6732	AlaGlyGlnThr-----ValLeuGluSerPheThrValSerSerSerAasp	6746
QY	3526	TCTCTCAAGTGTAAACGCCTAAATGTAAACACTCTTTTAAAGGAGACGATGTTTACTTG	3585
DB	6747	GlyThrGlnAlaGlnVal----SerValThrIleThrGlySerGlnAspAsnAlaValIle	6765
QY	3586	AATGGAGACTGCGCTTTTGTCAATGCTATGCGAGAGCTGAAGAAGTTCGATTATCTCA	3645
DB	6766	ThrGlyGluAlaIle-----GlyGlnValSerGlu	6775
QY	3646	GCTAATGGCGACAATTTAAACGATTACC GGACAAAACCATACATTATCATTTTACAGATTCT	3705
DB	6776	ThrAsnAaspSerGlnSerMetAlaThrGlyLysLeuAsnIleSerAaspLeuAaspSer	6795
QY	3706	CAA-----GGGCCA-----	3714
DB	6796	GlnGlnAlaHisPheValGlyGlyProGlnAlaGlyGlnTyrglyAlaPheThrLeuGln	6815
QY	3715	-----GTTCTTCAAAATVATGCCCTTCATTTCAGCAGCA-----	3747
DB	6816	GlyAaspGlyAsnTrpThrTyValLeuAaspAsnHisGlyValAlaMetAlaGlyLeuGly	6835
QY	3748	-----GAGACACTTACTCTGAGAGATTTTTCGACT-----	3777
DB	6836	ThrGlyAaspValValThrGluThrPheThrValArgSerLeuAaspAsnThrAlaGluThr	6855
QY	3777	-----	3777
DB	6856	ThrValThrIleThrIleAsnGlyLeuAaspGluProThrAlaGluGluIleGluAla	6875
QY	3778	-----CTGATGTTCTCGAAAAATGTTTCT	3801
DB	6876	AlaArgLeuAlaGluGlnArgLeuGlnGlnGluLeuGlnAlaSerAsnAaspLeuLeu	6895
QY	3802	TGCGGAGAAAAGGAATGATCTCCGGGAAAAACCGTGAGTATTTCCGAGCAGCGCAAGTG	3861
DB	6896	GlnGlyAlaIleGlyAlaGluGlyGly-----	6904
QY	3862	ATTTTCTGGGATACTCCGCGGGGTATTCCTTTATCTACTGTGCCAACCTCATCATCA	3921
DB	6905	-----AsnAlaGlyGlyThrAlaPro-----	6911
QY	3922	ACTCGCGCTGCTCCAACAGTTAGTATGATCTCGGAAAGGTCATATTTTCTGTAGAGACT	3981
DB	6912	ProProAasnGlyAsnProAlaAaspGlyAlaProGlyAaspLeu-----	6926
QY	3982	AGTTTGGAGATCTCAGGCGTCAAAAAGGGGTGATGTTCCGATAATAATCCGCGGAATTTTC	4041
DB	6927	---LeuGlyAlaGlyGlyAlaAlaIleGlyAlaThrGlyGlyThrThrGly-----	6943
QY	4042	CGAACAGTTTTTTCGAGGTGAAGAATAATAATGCTGTGTGTGAGCGCAGTGGTTCGCG	4101
DB	6944	GlyGlyThrThrAlaGlyThrGlyGlyGlyLeuGlyGlyGlyGly-PheGlySerGly	6963
QY	4102	TACACCATCAAGTAGC-----	4131
DB	6963	yAlaProAsnThrSerAlaGlyGlyLeuThrGlyGlyPheGlyGlyThrAlaThrGlyGly	6983
QY	4132	AAACTGTAAGGGAAGTTTCTTTTCACAGATAACGTAGCTCTTCGCGCA-----	4180
DB	6983	yThrValThrGlyGlyThrGlyGlyThrGlyGlyAlaAlaThrGlyGlyThrAlaThrGly	7003
QY	4180	-----	4180
DB	7003	yGlyThrAlaThrGlyGlyThrGlyGlyThrGlyGlyThrThrThrGlyGlyThrAlaThr	7023

```

4181  -GCGGAGTGGTTTATAAGGCATTGTGCTTTTCAAGACAACTGAAGGAGCATATTCTT 4239
      |||||
7023  rGlyGlyThrAlaThrGlyGlyThrGlyGlyAlaAlaGlyGlyThrGlyGlyAlaAla 7043
      |||||
4240  CCGAGGGAACACAGCATACGATGATTTAAGGATTTCTGCTGCTACTAATCAGGATCAGAA 4299
      |||||
7043  aGlyGly-----AlaAlaThrGlyGlyAlaGlyGly 7053
      |||||
4300  TACGAGACAGGAGCGGTGAGAGATTATTGCTCTCCAGATGATTCTGTAAGTTTGA 4359
      |||||
7053  yThrAlaThrGlyGlyThrGlyGlyAlaAlaThrGlyGlyThrAlaThrGly 7073
      |||||
4360  AGGCAATAAAGGTTCTATTGTTTGTGATTACAACTTTGCAAAAGCGAGCGCGAAGC 4417
      |||||
7073  yGlyThrGlyGlyAla-----AlaGlyGlyThrGlyGlyAla 7085

RESULT 31
US-11-052-554A-282
; Sequence 282, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 282
; LENGTH: 1345
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-282

```

QY 664 ---GGACACTCGTTGACCTTCGAGACATACCGACTTCTACA----- 702
 Db 120 AspSerGlnProValThrPheValAlaAspLysThrSerAlaLeuValValLeuGlnIle 139
 QY 703 -----AATGGG-----GCAGCTCTAAGTAATAGCGCTGCT 732
 Db 140 SerLysAsnGluIleThrGlyAsnGlyValAspSerAlaThrLeuThrAlaThrValLys 159
 QY 733 GATGACTGTTTACTATTGAGGGTTTTAAAGAAATATCTCTTTCCAAATTCGAATTCATTA 792
 Db 160 AspGlnPheAspAsnGluValAsnAsnLeuProValThrPheSerThrAlaSerSergly 179
 QY 793 CTGCGCTAGTCGCTGCT---GCAACGACTAATAAG----- 825
 Db 180 LeuThrLeuThrProGlyGluSerAsnThrAsnGluSerGlyIleAlaGlnAlaThrLeu 199
 QY 826 -----GGTAGCCAGACTCCGACGACACATCTACACCGCTCTAATGGTACT 870
 Db 200 AlaGlyValAlaPheGlyGluGlnThr---ValThrAlaSerLeuAlaAsnAsnGlyAla 218
 QY 871 ATTTATCTTAAACAGATCTTTTGTGTAATCAATAATAGAAAGTTCTCATTTATAGTAAT 930
 Db 219 SerAspAsnLysThrValHisPheIle----- 227
 QY 931 TAGTCTCTGGAGATGGGGAGCTATAGATGCTAGAGCTTAACGGTTCAAGGAATTAGC 990
 Db 228 -----GlyAspThrAlaAlaAlaLysIleLeuLeuThrProValProAspSer 244
 QY 991 AAGCTTTGTCTCTCAAGMAAACTGCTCAAGCTGATGGGGAGCTGTGCAAGTAGTC 1050
 Db 245 IleIleAla-----GlyThrProGlnAsnSerSerglySer-----ValIle 258
 QY 1051 ACCAGTCTCTGCTATGCTTAACAGGCTCTATTGCTTTGTA----- 1095
 Db 259 ThrAlaThrValValAspAsnAsnGlyPheProValLysGlyValThrValAsnPheThr 278
 QY 1096 -----GCGAATGTTGCGAGGATAGAGGGGGAGGATTCGCTGTTTCAGGATGGGCAG 1149
 Db 279 SerAsnAlaAlaThrAlaGluMetThrAsnGlyGlyGlnAlaValThrAsnGluGlnGly 298
 QY 1150 CAGGAGGTGTCATCATCTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACT 1209
 Db 299 LysAlaThrValThrTyThrAsnThrArgSerSerglyGluSerGlyAlaArgProAsp 318
 QY 1210 CGGTAGATGTTGATGGGAACGTAGCCCGAGTAGGAGGAGGATTTATCTCTACGGGAAC 1269
 Db 319 ThrValGlu----- 321
 QY 1270 GTTGCTTTCTGTAATAATGGAAAA-----ACCTGTTTCTCAACAATGTT 1314
 Db 322 ---AlaSerLeuGluAsnGlySerSerThrLeuSerThrSerIleAsnValAsnAlaAsp 340
 QY 1315 GCTTCTCTGTTTACATGCTGCTCTAAGCAACCA-----ACAAGTGGACAG 1359
 Db 341 AlaSerThrAlaHisLeuThrLeuLeuGlnAlaLeuPheAspThrValSerAlaGlyAsp 360
 QY 1360 GCTTCTTAAT-----ACGAGTAATAATTACGAGATGGAGGA----- 1395
 Db 361 ThrThrAsnLeuTyrIleGluValLysAspAsnTyrglyAsnGlyValProGlnGlnGlu 380
 QY 1396 -----GCTATCTTCTGTAAGATGTTGGCGACGCGGATCCAAATCACTCTGATCA 1446
 Db 381 ValThrLeuSerValSerProSerGluGlyValThrProSerAsnAsnAlaIleTyThr 400
 QY 1447 GTTCTCTTGGAGAGGGAGTAGTTTCTTTTAGCAATGTAGCTGCTGGGAAAGGG 1506
 Db 401 ThrAsnHisAspGlyAsn-----PheTyralaSerPheThrAlaThrLysAla 416
 QY 1507 GGAGCTATTATGCCAAAAGCTCTGGTGTCTAACTGTGGCCCTGTACAA----- 1557
 Db 417 Gly---ValTyGlnValThrAlaThrLeuGluAsnGlyAspSerMetGlnGlnThrVal 435
 QY 1558 ---TTTTAAGGAATATCCCTAATGATGTTGGGCGGATTTATTTAGGAGAAATCTGGAGAG 1614

Db 436 ThrTyrrValProAsnValAlaAsn----- 445
 QY 1615 CTCAGTTTATCTGCTGATATGGAGATATTATTTCATGATGGAAATCTTTAAAGAACAGCC 1674
 Db 446 IleSerLeuAlaAlaSerLysAspProValIleAlaAsnAsnAsnLeuThrThrLeu 465
 QY 1675 AAAGAGATGCTCGCATGTTAATGGGTAACTGCTGCTCTCACAGCCATT----- 1725
 Db 466 ThrAlaThrValAlaAspThrGluGlyAsnAlaIleAlaAsnSerGluValThrPheThr 485
 QY 1726 -----TCGATGGATCGGAGCGGAAATAAACACACATTA 1758
 Db 486 LeuProGluAspValArgAlaAsnPheThrLeuGlyAspGlyGlyValValThr----- 504
 QY 1759 AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
 Db 504 ----- 504
 QY 1819 AACCAGCCAGCGCAGCTCTCCAAACTTCTAAAAATTAAACGATGGTGAAGGATACACAGGG 1878
 Db 505 AspThrGluGlyLysAlaLysValThrLeuLysGlyThrLysAlaGlyAlaHisThrVal 524
 QY 1879 GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGACAGCA 1938
 Db 525 ThrAlaSerMetAlaGlyLysSer----- 533
 QY 1939 AGGATTGTTCTCGTGAAGCGCAAAATTAATCATAGTGAATCTCTAAGTCAGACAGGTGG 1998
 Db 534 -----GluGlnLeuValValAsnPheIleAlaAspThr----- 544
 QY 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACACCACCAACAG 2058
 Db 544 ----- 544
 QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTCCAATCTGCAATTTGCTCTTCTTCTTTG 2118
 Db 545 -----LeuThrAlaGlnValAsnLeuValThrGluAspAsnPhe 558
 QY 2119 TTAGCAACAAT-----GCAGTTACGAATCTCTTACCAAT 2154
 Db 559 IleAlaAsnAsnValGlyMetThrArgLeuGlnAlaThrValThrAspGlyAsnGlyAsn 578
 QY 2155 CTTCAGCGCAAGATTTCTATCTGTCAGTCACTGTCAGCACACTGCTGTTCTGTTACA 2214
 Db 579 ProLeuAlaAsnGlu-----AlaValThr 586
 QY 2215 ATTAGTGGGCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGTATGATTGG 2274
 Db 587 PheThr----- 588
 QY 2275 CTAGGTTCTAATCAAAAATCAATGTCCTGAAATTTACAGTTAGGAGCTAAGCCCCAGCT 2334
 Db 589 -----LeuProAla 591
 QY 2335 AATGCCCATCAGATTTGACTCTAGGAAATCAGATGCCCTAAGTATGGCTATCAAGGAAGC 2394
 Db 592 AspValSerAlaSerPheThrLeuGly-----GlnGlyGly 603
 QY 2395 TGGAGCTTGGCTGGATCCTTAATACAGCAAAATATGCTTATCTCTGAAAGCTACA 2454
 Db 604 SerAlaIle-----ThrAspIleAsnGlyLysAlaGluValThrLeuSer 618
 QY 2455 TGGACTAAAACCTGGG-----TATAATCTCGGGCT 2484
 Db 619 GlyThrLysSerGlyThrTyrrProValThrValSerValAsnAsnTyrrGlyValSerAsp 638
 QY 2485 GAGCAGTAGCTCTTGGTTCCTCAAAATAGTTATGGGGATCCATTTTATAGATATCACT 2544
 Db 639 ThrLysGlnValThrLeuIleAlaAspAlaGlyThrAlaLysLeuAlaSerLeuThrSer 658
 QY 2545 GCGCATTCAGCAATTCAGAGCAAGTGTGATGGGCGCTCTTATTGTCGAGGATTATGGGTT 2604

Db 659 ValTyrSerPheValValSerThrThrGluGlyAlaThrMetThrAlaSerVal----- 676
Qy 2605 TCTGGAGTTTCGAATTTCTTCTATCATGACCGCGATGCTTTAGGT-----CAGGGA 2655
Db 677 -----ThrAspAlaAsnGlyAsnProValGluGly 686
Qy 2656 TATCGGTATATTAGTGGGGCTTATTCCTTAGGAGCAAACTCC----- 2697
Db 687 IleTyrValAsnPheArgGlyThrSerValThrLeuSerSerThrSerValGluThrAsp 706
Qy 2698 -----TACTTTGGATCATCGATGTTTGGTCTAGCATTT----- 2730
Db 707 AspArgGlyPheAlaGluIleLeuValThrSerThrGluValGlyLeuTyrValSer 726
Qy 2731 -----ACCGAAGTATTGGTAGTCTAAGATTATGTAGTGCT 2769
Db 727 AlaSerLeuAlaAspLysProThrGluValIleSerArg----- 739
Qy 2770 CGTTCCATCATCATGCTTGGATGATCGGTTTATCTATCTATCTACCAACAAGCTTTATGT 2829
Db 740 LeuLeuAsnAlaLysAlaAspIleAsnSerAlaThrIleThrSerLeuGluIleProGlu 759
Qy 2830 GATCTCTATTGTCGGA---GATGCGTTTATCCGTCGTAGCTACGGG-----TTTGGG 2880
Db 760 GlyGlnValMetValAlaGlnAspValAlaValLysAlaHisValAsnAspGlnPheGly 779
Qy 2881 AATCAGCATATGAAACCTCATATACATTATGTCAGAGGAG-----AGCGATGTCGTTGG 2934
Db 780 AsnProIleLeuAsnGluSerValThrPheSerAlaGluProProGluHisMetThrIle 799
Qy 2935 GATAATACTGTCGTCGAGAGATGGCGGGATTACCGATGTCGATTAATCCATCT 2994
Db 800 SerGlnAsnIleValSerThrAspThrHisGlyIleAlaGluValThrMetThrPro--- 818
Qy 2995 AAGCTCTATTGTAATGAGTTGCGCTCTTCGTCGAAGCTGAGTTCCTTATGTCGCGATCAT 3054
Db 819 -----GluArgAsnGlySerTyrMetValLysAlaSerLeuAlaAsnGly 833
Qy 3055 GAATCTTTTACAGAGAA-----GGCGATCAAGCTCGGCGATTCAAGAGCGGA 3102
Db 834 SerSerTyrGluLysAspLeuValIleAspGlnLysLeuThrLeuSerAlaSerSer 853
Qy 3103 CATCTCTAATATCATGTTCTCTGTTGGAGTGAAGTTTCAT-----CGATGTCCT 3153
Db 854 ProLeuIleGlyValAsnSerProThrGlyAlaThrLeuThrAlaThrLeuThrSerAla 873
Qy 3154 AGTACACATCTTAATAATAGCTTTATGCGCGCTTATATCTGTCGATGCTTATCGCACC 3213
Db 874 AsnGlyThrProValGluGlyGlnValIleAsnPheSerValThrProGluGlyAlaThr 893
Qy 3214 ATCTCTGTTACTGAGACAAGCTCTTATCCCATCAAGAG----- 3252
Db 894 LeuSerGlyGlyLysValAlaGthrAsnSerSerGlyGlnAlaProValValLeuThrSer 913
Qy 3253 -----ACATGGACA---ACAGATGCCCTTTCATTAGCAACATGAGGTGTG 3297
Db 914 AsnLysValGlyThrTyrThrValThrAlaSerPheHis-----AsnGlyValThr 930
Qy 3298 GTTAGAGGATCATGTATGCTTCTCAACAAGTAATATAGAATATATGTCATGCAAGGA 3357
Db 931 IleGlnThrGlnThrIleValLysValThrGlyAsnSerSerThr---AlaHisValAla 949
Qy 3358 TATGAGTATCGAGATGCTTCTCGAGGCTATGGT-----TTGAGT 3396
Db 950 SerPheIleAlaAspProSerThrIleAlaAlaThrAsnSerAspLeuSerThrLeuLys 969
Qy 3397 GCAGGAAGTAGAGTCGGGTTCTTAAATATTTGGTTAGTAGTATTAGTTAGCATGCCCT 3456
Db 970 Ala-ThrValGluAspGlySerGlyAsnLeu---IleGluGly-----LeuThrValTyr 986
Qy 3457 TTTTCTTTGAGATCTACATCATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3510
Db 987 PheAlaLeuLysSerGlySer-----AlaThrLeuThrSerLeuThrAlaVal 1002

Qy 3511 -----TATGATTTCGCGAGCTCTCT-----CAAGTGTTTAACGCCCTAATGTAAACC 3555
Db 1003 ThrAspGlnAsnGlyIleAlaThrThrSerValArgGlyAlaIleThrGlySerValThr 1022
Qy 3556 ACTCCTTTTAAAGGAGACGATGTTTACTTGAATGGAGACTGGCTTTTGTCAATGTC--- 3612
Db 1023 Val-----SerAlaValThrThrAlaGlyMetGlnThrValAspIleThr 1038
Qy 3613 ---TATCAGCAGCTGAAGAAGGT-----TCGATTATCTCAGCTAATGCGCACAAATTA 3663
Db 1039 LeuValAlaGlyProAlaAspAlaSerGlnSerValLeuLysAsnAsnAspSerSerLeu 1058
Qy 3664 -----ACGATTACCGGACAAACCATCATATTATCATTTACAGATTCTCAAGGG 3711
Db 1059 LysGlyAspPheThrAspSerAlaGluLeuHisLeuValLeuHisAspIleSerGlyAsn 1078
Qy 3712 CCA-----GTTCTTCAAAATATCCCTTCATTTTCAGCAGGAGACACTTACTCTGAGA 3765
Db 1079 ProIleLysValSerGluGlyLeuGluPheValGlnSerGlyThrAsnAlaProTyrVal 1098
Qy 3766 GATTTTTCGAGTCTGATGTTCTCGAAATGTTTCTGCGGAGAGAAAGGAAATGATCTCC 3825
Db 1099 GlnValSerAlaIleAspTyrSerLysAsnPheSerGlyGluTyrLysAlaThrValThr 1118
Qy 3826 GGGAAAAACCGTGAGTATTTCGCGAGCAGCGAAGTATTTCTGGGATTAACCTCCGTGGG 3885
Db 1119 GlyGlyGlyGluGlyIleAlaThrLeuIleProValLeu-----AsnGlyValHis 1135
Qy 3886 TATTCCTCTTATCTACTGTGCAACCTCATCATCACTCCGCTGCTCCACAGTTAGT 3945
Db 1136 GlnAlaGlyLeuSerThr-----ThrIleGln 1144
Qy 3946 GATCTCGAAAGGGTCTATTTTCTGTAGAGACTAGTTTGTGAGATCTCAGCGCTCAAA 4005
Db 1145 PheThrArgAlaGluAspLysIleMetSerGlyThrValLeu---ValAsnGlyAlaAsn 1163
Qy 4006 AAAGGGGTCAATGTCGTAATAATGCGGGAATTTTCGGAACACAGTTTTTTCAGGTAAGAAT 4065
Db 1164 LeuProThrThrThrPheProSerGlnGlyPheThrGlyAlaTyrTyrGlnLeuAsnAsn 1183
Qy 4066 AATAATAATGCT---GGTGGTGGAGCAGCTGGTTCGCTACACCATCAAGTACGACTTT 4122
Db 1184 AspAsnPheAlaProGlyLysThrAlaAspTyrGluPheSer-SerSerAlaSerTr 1203
Qy 4123 TACAGTTAAAACTGTAAAGGGAAGCTTTCTTTCACAGATAAGCTAGCTCT 4174
Db 1203 pValAspValAspAlaThrGlyLysValThrPhe---LysAsnValGlySer 1219

RESULT 32

US-11-052-554A-280
; Sequence 280, Application US/11052554A
; Publication No. US2005028886A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 280
; LENGTH: 1588
; TYPE: PRN
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-280
Alignment Scores:

Pred. No.:	4, 69e-09	Length:	1588
Score:	205.00	Matches:	293
Percent Similarity:	30.9%	Conservative:	203
Best Local Similarity:	18.3%	Mismatches:	577
Query Match:	2.6%	Indels:	531
DB:	11	Gaps:	69
US-10-701-844-1 (1-4435) x US-11-052-554A-280 (1-1588)			
QY	523	TATACCTGTTATAGGAGATCCGATGGGACTACTGTTTTTCTCGCAGGAGATTAAACATTA	582
DB	16	TyrThrValThrSerGluThrAlaLysSerArgGlyLysSerGlyArgSerGlySerLysLeu	35
QY	583	AAAAATCTTGACAAATCTATTGCGACCTTGGCTTTAAAGTTGTTTGGGAACCTATTAGGG	642
DB	36	-----LeuIleSerAlaLeuValAla	47
QY	643	AGTTTTACTGTTTATAGGAGGAGGACACTCGTTGACTTTTCGAGAACATACGGACTTCTACA	702
DB	48	SerPheGlyAlaLeuAlaAsnAla	58
QY	703	AATGGGCGAGCTCTA-----AGTAATAGCGCTGCTGATGGACTGTTTACTATT---	750
DB	59	AsnGlyGlnGlyValAspTyrGlySerGlySerAlaGlyAspGlyTyrValAlaIleGly	78
QY	751	GAGGTTTTAAAGAAATATCTTTTCCAAATTCGCAATTCATTTACTTGGCGTACTGCTGCT	810
DB	79	LysGlyAlaLysAlaAsnThrPheMetAsn	88
QY	811	GCAACGACTAATAAGGGTAGCCAGACTCCGACGACACATCTACACCGCTCTAATGGT---	867
DB	89	-----ThrSerGlySerSerThrAlaValGlyTyrAspAlaIleAlaGluGlyGln	105
QY	868	-----ACTATTTATTTCTAAACAGATCTTTTGTACTCAATAATGAGAAGTTCTCA	918
DB	106	TyrSerSerAlaIleGlySerLysThrHis-----AlaIleGlyGlyAlaSerMetAla	123
QY	919	TTCTATAGTAATTTAGTCTCT---GGAGATGGGGA---GCTATAGATGCTTAAGAGCTTA	972
DB	124	PheGlyValSerAlaIleSerGluGlyAspArgSerIleAlaLeuGlyAlaSerSerTyr	143
QY	973	ACGGTT-----CAAGGAATTAGCAAGCTTTGT	999
DB	144	SerLeuGlyGlnTyrSerMetAlaLeuGlyArgTyrSerLysAlaLeuGlyLysLeuSer	163
QY	1000	GTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACACAGTTTC	1059
DB	164	IleAlaMetGlyAspSerSerLysAlaGlu---GlyAlaAsnAlaIleAlaLeuGlyAsn	182
QY	1060	TCTGCTATGGCTAACGAGCTCCTATTGCTTTTGTAGGAAATGTTGCGAGGAGTAAGA---	1116
DB	183	AlaThrLysAlaThrGluIleMetSerIleAlaLeuGlyAspThrAlaAsnAlaSerLys	202
QY	1117	-----GGGGGGGATGCTGCTGTTCCAGATGGCGGAGGAGTG	1158
DB	203	AlaTyrSerMetAlaLeuGlyAlaSerSerValAlaSerGluGluAsnAlaIleAlaIle	222
QY	1159	TCACTACTCTTCAACGAGATCCAGTAGTAAGTTTTCAGAAATACTCGGTAGAG	1218
DB	223	GlyAlaGluThrGluAlaAlaGluAsnAlaThrAlaIleGlyAsnAsnAlaLysAlaLys	242
QY	1219	TTTGATGGGAACGTAGCCGAGTAGAGGAGGATTTACTCTCAGGAACTTCGCTTC	1278
DB	243	GlyThrAsnSerMetAlaMetGlyPheGlySerLeuAlaAspLysValAsnThrIleAla	262
QY	1279	CTGAATAATGGAAACCTTGTCTCAACAATGTTGCTTCTCCTGTTTACATTCGCTGCT	1338
DB	263	LeuGlyAsnGlySerGlnAlaLeuAlaAspAsnAla-----IleAlaIle	277
QY	1339	AAGCAACCAACGAGGAGGAGCTCTTAATACGAGTAATAATTACGAGATCGA---	1395
DB	278	GlyGln-----GlyAsnLysAlaAspGlyValAsp	287

QY	1396	GCTATCTTCTGAAGAATGGTGGCAAGCA-----	1425
DB	288	AlaIleAlaLeuGlyAsnGlySerGlnSerArgGlyLeuAsnThrIleAlaLeuGlyThr	307
QY	1426	-----GGATCCCAATAACTCT---	1440
DB	308	AlaSerAsnAlaThrGlyAspLysSerLeuAlaLeuGlySerAsnSerSerAlaAsnGly	327
QY	1441	---GGATCAGTTTCTTTGATGGAGGAGGAGTAGTTCTTTCTTAGTACGATAGTCTGCT	1497
DB	328	IleAsnSerValAlaLeuGlyAlaAspSerIleAlaAspLeuAspAsnThrValSerVal	347
QY	1498	GGG-----AAAGGGGAGCTATT-	1515
DB	348	GlyAsnSerSerLeuLysArgLysIleValAsnValLysAsnGlyAlaIleGlySerAsp	367
QY	1516	-----TATGCCAAAAAGCTCTCGTTGCTTAAC---	1542
DB	368	SerTyrAspAlaIleAsnGlySerGlnLeuTyrAlaIleSerAspSerValAlaLysArg	387
QY	1543	-----TGTGGCCCTGTACAAATT	1560
DB	388	LeuGlyGlyAlaAlaValAspValAspGlyThrValThrAlaProThrTyrAsn	407
QY	1561	TTAAGGATATCGCT-----AATGATGGTGGAGGATTTATTATTAGGAGATCTGGAGAG	1614
DB	408	LeuLysAsnGlySerLysAsnAsnValGlyAlaAlaLeuAlaValLeuAspLysAsnThr	427
QY	1615	CTC-----	1617
DB	428	LeuGlnTyrAspGlnThrLysGlyLysTyrSerAlaAlaHisGlyThrSerSerProThr	447
QY	1618	-----AGTTTATCTGCTGATTTATGGAGATATTATT	1647
DB	448	AlaSerValIleThrAspValAlaAspGlyThrIleSerAlaSerSerLysAspAlaVal	467
QY	1648	TTCGATGGATCTTAAAGAACACGCAAGAGATGCTGCGATGTTTAATGGCTAACT	1707
DB	468	AsnGlySerGlnLeuLysAlaThrAsnAspAspValGluAlaAsnThrAlaAsnIleAla	487
QY	1708	GTGCTCTCACAAGCATTTTCGATGGATCGGAGGG-----AAAAATA	1749
DB	488	ThrAsnThrSerAsnIleAlaThrAsnThrAlaAsnIleAlaThrAsnThrThrAsnIle	507
QY	1750	ACGACATTAGAGCTTAAAGCAGGG-----CATCAGATTCTCTTTAATGATCCC	1797
DB	508	ThrAsnLeuThrAspSerValGlyAspLeuGlnAlaAspAlaLeuLeuTyrAsnGluThr	527
QY	1798	ATCGAGATGGCAACGGAAATAACCGCCAGCGAGTCTCCAACTCTTAAATAATTAAC	1857
DB	528	LysLysAlaPheSerAlaAlaHisGlyGlnAspThrThrSerLysIleThrAsnValLys	547
QY	1858	GATGGTGA-----GGATACACA	1875
DB	548	AspAlaAspLeuThrAlaAspSerThrAspAlaValAsnGlySerGlnLeuLysThrThr	567
QY	1876	GGGGATATTGTT-----TTTGCTAATGGAAGCAGTACTTTGTACCAA	1917
DB	568	AsnAspAlaValAlaThrAsnThrThrAsnIleAlaAsnAsnThrSerAsnIleAlaThr	587
QY	1918	AATGTTACGATAGACCAAGGAAGGATTTCTTCTGTAAGGCAAAATATCATCTGAAT	1977
DB	588	AsnThrThr-----AsnIleSer	593
QY	1978	TCTCTAAGTCAGACAGGTGGGAGTCTGTATATATGGAAGCT-----	2016
DB	594	AsnLeuThrGluThrValThrAsnLeuGlyGluAspAlaLeuLysTyrTipAspLysAspAsn	613
QY	2016	-----	2016
DB	614	GlyValPheThrAlaAlaHisGlyThrGluThrThrSerLysIleThrAsnValLysAsp	633
QY	2017	-----GGGAGTACATGGGATTTGTAACTCCCAACACCACCAACAGCT	2061

QY 3874 AACTCCGGGGTATTCTCCCTTATCTACTGTGCCAACCTCATCATCACTCCGCTGCT 3933
 Db 1249 AenSerIleAlaIleGlyAenGlySerThrThrArgGlyAlaGlnThrAenThr 1268
 QY 3934 CCAACAGTTAGTGTCCGGAAGGGTCTATT-----TTTTCTGTGTA-----CAG 3978
 Db 1269 AlaTyAenMetAspAlaProGlnAenSerValGlyGluPheSerValGlySerAlaAsp 1288
 QY 3979 ACTAGTTTGGAGATCTCAGGGCTCAAAAAGGGGTGTCATGTTTCGATAATAATGCCGGGAAT 4038
 Db 1289 GlyGlnArgGlnIleThrAenValAlaAlaGly---SerAlaAspThrAspAlaValAen 1307
 QY 4039 TTCGGAACAGTTTTCGAGGTAAAGTAATAATAATGCTGGTGGAGGAGTGGGTTTC 4098
 Db 1308 ValGly-GlnLeu-----LysValThrAs 1315
 QY 4099 CGCTACACCACTACAGTCTTTTACAGTTTAAACCTGTAAGGGAAGTTCTTTCAC 4158
 Db 1315 pAlaGlnValSerGlnAenThrGlnSerIleThrAenLeuAenAspAenArgValThrAenLe 1335
 QY 4159 AGAT-----AACGTAGCTCTTTCGAGGCGGAGTGGTTTATAAGGCATTGT 4206
 Db 1335 uAspSerArgValThrAenIleGluAenGlyIleGlyAspIleValThrGlySerTh 1355
 QY 4207 G---CTTTTCAAGCAATGAAGGAGGCAATATTTCCGAGGGAACACAGCATACGATGA 4263
 Db 1355 rlySerThrPheThrAenThrAspGlyVal----- 1365
 QY 4264 TTTAAGGATCTTGTCTACTTAATCAGATCAGATCAGATCAGGAGCAGGCGGNGAGG 4323
 Db 1366 -----AspAlaSerAlaGlnGlyAspSer---ValAlaIleGlySerGI 1380
 QY 4324 AGTTATTGCTCTCCAGATGATTCGTAAAGTTTGAAGCAATAAGGTTCTATTGTTTT 4383
 Db 1380 ySerIleAlaAlaAlaAspAenSerValAlaLeu-----GlyThrGlySerValAlaTh 1398
 QY 4384 TGATTACAC 4393
 Db 1398 rGluGluAen 1401

RESULT 33

US-11-052-554A-92
 ; Sequence 92, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE

; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A

; PRIORITY FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227

; PRIOR FILING DATE: 2004-07-20

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004

; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 92

; LENGTH: 2399

; TYPE: PRT

; ORGANISM: Helicobacter pylori J99

US-11-052-554A-92

Alignment Scores:
 Pred. No.: 8,46e-09 Length: 2399
 Score: 203.00 Matches: 218
 Percent Similarity: 29.9% Conservative: 139
 Best Local Similarity: 18.3% Mismatches: 340
 Query Match: 2.6% Indels: 496
 DB: 11 Gaps: 52

US-10-701-844-1 (1-4435) x US-11-052-554A-92 (1-2399)

QY 469 ATCATGTTCTCTCAAGGAATTTACGAT-----GGGAGACGTTAACTGTATCA 516
 Db 1629 LeuLeuSerProArgGlyLeuHisAspPheTrpGlnLysGlyTyPheAenPheLeuSer 1648
 QY 517 TTTCCCTATATCTGTTATAGGAGATCCGAGT-----GGGACTACTGTT 558
 Db 1649 AsnGlyTyPheValPheValAenAenSerSerPheSerAenAlaThrGlyGlySerLeuAen 1668
 QY 559 TTTTCTCGAGGAGTTAAACATTAATAAATCTTCACAAATCTTATATGACGCTTTGCTTTTA 618
 Db 1669 PheValAlaAenLysSerIlePheAenGlyAspAenThrIleAsp----- 1684
 QY 619 AGTTGTTTGGAACTTATAGGAGCTTTACTGTTTATAGGAGGAGACACTCGTTGACT 678
 Db 1685 -----PheSerLysTyPheGlnGlyAlaLeuIlePheAlaSerAenGlyValSerAenIle 1702
 QY 679 TTCGAGAACATACGACT-----TCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCT 732
 Db 1703 -----AenIleThrThrLeuAenAlaThrAenGlyLeuSerLeu---AsnAlaGlyLeu 1719
 QY 733 GATGACTGTTTACTATTAGGGTTTAAAGAAATATCTCTTTCGAAATTCGCAATTCATTA 792
 Db 1720 AenAenValSerValGlnLysGlyGluIleCysIleAenLeuAlaAenCys----- 1736
 QY 793 CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 849
 Db 1737 -----ProThrThrLysAen 1741
 QY 850 TCTACACGCTCTAATGGTACTATTATTATCTAAACAGATCTTTGTTACTCAATTAATGAG 909
 Db 1742 SerSerProAlaAenSerSerValThrProThrAenGluSerLeuSerValHisAlaAen 1761
 QY 910 AAGTCTCATCTTATAGTAATTTAGTCTCTCGAGATGGGAGCTATAGTCTTAAGAGC 969
 Db 1762 AenPheThrPheLeuGlyThrIleIleSer-----AenGlyAlaIleAspLeuSerGln 1779
 QY 970 TTAACGGTTCAGGAATTAGCAAGCTTTGTGTCTTCCAGAAAT---ACTGCTCAAGCT 1026
 Db 1780 ValThrAenAenSerValIleGlyThrLeuAenLeuAenGluAenAlaThrLeuGlnAla 1799
 QY 1027 GATGGGGAGCTTGTCAAGTAGTACACAGTCTTCTGCTATGGCTAAAGAGGCTCCTATT 1086
 Db 1800 AenAen-----LeuThrIleThrAenAlaPheAenAen----- 1810
 QY 1087 GCCTTTGTAGCGAATGTTGCAGGAGTAAGAGGGGAGGATGCTGCTTCAGGATGG 1146
 Db 1810 ----- 1810
 QY 1147 CAGCAGGAGTGTGATCATCTACTTCAACAGAGATCCAGTAGTAGTATTTTCCAGAAAT 1206
 Db 1811 -----AlaSerAenSerThrAlaAenIleAspGlyAenPheThrLeuAenGln 1827
 QY 1207 ACTCGGTAGAGTTTGTATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTCAAGGG 1266
 Db 1828 AlaThrLeuSerThrAenAlaSerGlyLeuAenValMetGlyAenPheAenSerTyGly 1847
 QY 1267 AACGTTGCTTCTCGAATAATAGAAAAACCTTGTCTTCTCAAAATGTTGCTTCTCTGTT 1326
 Db 1848 AspLeuValPhe-----AenLeuSerHisSerVal 1857
 QY 1327 TACATTGCTTAAGCAACCAAGTGGACAGGCTTCTTAATACGAGTAATTAATACGA 1386
 Db 1858 -----SerHisAlaIleAenThrGlnGlyThr----- 1867
 QY 1387 GATGGAGGAGCTATCTTCTGTAAGAAATGCTGGCAGCAGGATCCAATCACTCTGATCA 1446
 Db 1868 -----AlaThrIleMetAlaAenAenAenProLeu 1877
 QY 1447 GTTTCCTTTGATGGAGGAGGAGTAGTATTTCTTTAGTACAAATGTTAGCTGCTGGAGGG 1506
 Db 1878 IleGlnPheAen-----AlaSerSerLysGlu 1886


```

Db 2290 AlaaspProAsnLysLysSerValLeuThrIleAsnPheAlaLeuLysSerArgHis---- 2308
Qy 3574 GATGTTTACTTGAATGGAGACTGGCGCTTTTGTCAATGTCTATGCGAGGAGCTGAAGAAGGT 3633
Db 2309 -----TyrPheAsnLysAsn---SerTyrTyrPheValIleAlaAspValGlyArgAsp 2325
Qy 3634 TCGATTATCTCAGCTAATGGGACAAATTTAAGGATTACCGGACAAACCAATACATTATCA 3693
Db 2326 LeupheIleAsnSerMetGlyAspLysMetValArgPheIleGlyAsnAsnThrLeuSer 2345
Qy 3694 TTTACAGATTCTCAAGGGCCAGTCTTCAAAATTAATGCG---TTCAATTTCAAGCAGGAGAG 3750
Db 2346 TyrArgAsp-----GlyGlyArgTyrAsnThrPheAlaSerIleIleThrGlyGlyGlu 2363
Qy 3751 ACATCTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAATAATGTTTCTTGGCGAGAA 3810
Db 2364 -----IleArgLeuPheLysThr-----PheTyrValAsnAlaGlyIleGlyAla 2378
Qy 3811 AAGGAAATGATCTCCGGGAAACCGTGAGTATTTCGGGA 3849
Db 2379 ArgPheGlyLeuAspTyrLysAspIleAsnIleThrGly 2391

RESULT 34
US-11-052-554A-374
; Sequence 374, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 374
; LENGTH: 2015
; TYPE: PRT
; ORGANISM: Neisseria meningitidis Z2491
US-11-052-554A-374

Alignment Scores:
Pred. No.: 8, 64e-09 Length: 2015
Score: 202.50 Matches: 298
Percent Similarity: 32.7% Conservative: 201
Best Local Similarity: 19.5% Mismatches: 585
Query Match: 2.6% Indels: 442
DB: 11 Gaps: 68

US-10-701-844-1 (1-4435) x US-11-052-554A-374 (1-2015)
Qy 448 GGGGGGGGATTCGACGAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAGAGCTTGA 507
Db 284 GlyGlyMetTyrAlaAspSerIle----- 291
Qy 508 ACTGTATCATTTCCCTATCTCTTATAGGAGATCCGAGTGGGACTACTGTTTCTGCA 567
Db 292 -----ThrLeuIleAlaAsnGluLysGlyValGlyValLysAsnAla 305
Qy 568 GGAGAGTTAAACATTAATAAATCTTGACAATCTTATGACGCTTGTGCTTTAAAGTTGTTT 627
Db 306 GlyThrLeu-----GluAlaAlaLysGlnLeuIleValThrSerSer 319
Qy 628 GGGAACTATTAGGGAGT-----TTTACTGTTTATAGGAGAGACACTCGTTG 675
Db 320 GlyArgIleGluAsnSerGlyArgIleAlaThrAlaAspGlyThrGluAlaSerPro 339
Qy 676 ACTTTGAGAACTACCGACTCTTACAAATGGGCGAGCT-----CTAAGTAATAGC 726

```

```

Db 340 ThrTyrLeuSerIleGluThrThrGluLysGlyAlaAlaGlyThrPheIleSerAsnGly 359
Qy 727 GCT-----GCTGATGGACTGCTTTTACTATTGAGGGTTTTAAAGAATTATCTTTTCC 777
Db 360 GlyArgIleGluSerLysGlyLeuValIleGluThrGlyGluAspIleSerLeuArg 379
Qy 778 AATTGCAATTCAATTCTGCGTACTGCTGCTGCAACGACTAATAGGTTAGCCAGACT 837
Db 380 Asn-----GlyAlaValValGlnAsnAsnGlySerArgPro 391
Qy 838 CCGACGACACA-----TCTACACCGCTCTAATCGTACTATTATTCTTAAACAGATCTTTTG 894
Db 392 AlaThrThrValLeuAsnAlaGlyHisAsnLeuValIleGluSerLysThrAsn----- 409
Qy 895 TTACTCAATAATAGAGAGTTCTCATTTCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCT 954
Db 410 ---ValAsnAsnAlaLysGlySer-----AlaAsnLeuSerAla-----GlyGlyArg 424
Qy 955 ATAGATGCTAAGAGCTTAACGGTTTCAAGGAATTAGCAAGCTTTTGTGTTCTTCCAGAAAT 1014
Db 425 ThrThrIleAsnAspAlaThrIleGlnAlaGlySerSerValTyrSerSerThrLysGly 444
Qy 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGCTATGCTAAC 1074
Db 445 AspThrGluLeuGlyGluAsnThrArgIleIleAlaGluAsnValThrValLeuSerAsn 464
Qy 1075 GAGGCTCCTATTGCTTTGTAGCGAATGTTTGCAGAGTAGAGGGGGAGGAGTGTCTCT 1134
Db 465 GlySer-----IleGlySerAlaAlaValIleGluAlaLysAspThrAlaHis 480
Qy 1135 GTTCAGGATGGGCGAG-----CAGGAGGTGTCATCATCTACTTCAACA 1176
Db 481 IleGluSerGlyLysProLeuSerLeuGluThrSerThrValAlaSerAsnIleArgLeu 500
Qy 1177 GAAGATCCAGTAGTAAGTTTTTCCAGAAATACTCGGTAGAGTTTGTAGGAAACGTAGCC 1236
Db 501 AsnAsnGlyAsnIleLysGlyGlyLysGlnLeuAlaLeuAlaAspAsnIleThr 520
Qy 1237 CGAGTAGGAGGAGGATTTACTCTCAGGAAACGTTGCTTCTGTAATTAATGGAACCAACC 1296
Db 521 AlaLysThrThrAsnLeuAsnThrProGlyAsnLeu---TyrValHisThrGlyLysAsp 539
Qy 1297 TTGTTTCTCAAC-----AATGTTGCTTCTCTGTTTACATTCGCTGCTAAGCAACCAACA 1350
Db 540 LeuAsnLeuAsnValAspLysAspLeuSerAlaAlaSerIleHisLeuLysSerAspAsn 559
Qy 1351 AGTGACAGAGCTTCTAATACGAGTAATAATACGAGATGCGAGGAGCTATCTCTGTAAG 1410
Db 560 AlaAlaHisIleThrGlyThrSerLysThrLeuThrAlaSerLysAspMet----- 576
Qy 1411 AATGTTGCGCAAGCAGGA-----TCCAATAACTCTGGA 1443
Db 577 ---GlyValGluAlaGlyLeuLeuAsnValThrAsnThrAsnLeuArgThrAsnSerGly 595
Qy 1444 TCAGTTTCTTGTGAGGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTCTGGGAAA 1503
Db 596 AsnLeu-----HisIleGlnAlaLys 603
Qy 1504 GGGGGA---GCTATTATTATGCCAAAAGCTCTCGGTTGCTAACTGTGGCGCTGTACAAATTT 1560
Db 604 GlyAsnIleGlnLeuArgAsnThrLysLeuAsnAlaAlaLysAlaLeuGluThrThrAla 623
Qy 1561 TTAAGG---AATATCTGCTAATGATGTTGGAGCGAATTTATTTAGGAGAATCTCGAGAGCTC 1617
Db 624 LeuGlnGlyAsnIleValSerAsp---GlyLeuHisAlaValSerAlaAspGlyHisVal 642
Qy 1618 AGTTTATCTGCTGATTTAGGAGATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAA 1677
Db 643 SerLeuLeuAlaAsn---GlyAsnAlaAspPheThrGlyHisAsnThrLeuThrAlaLys 661
Qy 1678 GAGAATGCTGCGAGTAAATGCG----- 1701
Db 662 -----AlaAspValAsnAlaGlySerValGlyLysGlyArgLeuLysAlaAspAsn 678

```

QY 1702 -----GTAAGTGTCTCACAAGCCATTTTCGATG 1731
Db ThrAsnIleThrSerSerSerGlyAspIleThrLeuValAlaGlyAsnGlyIleGlnLeu 698
QY 1732 CGATCGGA-----GGGAAATAAGCAGATTAAAGAGCTAAAGCA 1770
Db GlyAspGlyLysGlnArgAsnSerIleAsnGlyLysHisIleSerIleLysAsnAsnGly 718
QY 1771 GGGCATCAGATTCTTTAATGATCCCATCGATGGCAACGGA----- 1815
Db GlyAsnAlaAspLeuLysAsnLeuAsnValHisAlaLysSerGlyAlaLeuAsnIleHis 738
QY 1816 -----AATAACCGACGCGCAGCTTCC----- 1839
Db SerAspArgAlaLeuSerIleGlnAsnThrLysLeuGluSerThrHisAsnThrHisLeu 758
QY 1840 -----AACTTCTAAAATAACGATGCTGAAGGATAC----- 1872
Db AsnAlaGlnHisGluArgValThrLeuAsnGlnValAspAlaTyAlaHisArgHisLeu 778
QY 1873 -----ACAGGGGATATTGTTTT----- 1890
Db SerIleThrGlySerGlnIleTrpGlnAsnAspLysLeuProSerAlaAsnLysLeuVal 798
QY 1891 CCTAATGGA-----AGCAGTACTTCTACCAAAATGTTACGATA 1929
Db AlaAsnGlyValLeuAlaLeuAsnAlaArgTySerGlnIleAlaAspAsnThrThrLeu 818
QY 1930 GAGCAAGGAAGGATTGTTCTTCGTGAAAGGCAAAATTA-----TCAGTGAAT 1977
Db ArgAlaGlyAlaIleAsnLeuThrAlaGlyThrAlaLeuValLysArgGlyAsnIleAsn 838
QY 1978 -----TCTCTAAGTCAGACAGGTGG 1998
Db TrpSerThrValSerThrLysThrLeuGluAspAsnAlaGluLeuLysProLeuAlaGly 858
QY 1999 AGTCGTATATGAGAGCTGGAGTACATGGGATTTGTAACTCCCAACCAACCAACAG 2058
Db ArgLeuAsnIleGluAlaGlySer-----GlyThrLeuThrIleGluPro 873
QY 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTCCCAATCTGCATTTGTCTCTTTCTTTG 2118
Db AlaAsnArgIleSerAlaHisThrAspLeuSerIleLysThrGlyGlyLys 890
QY 2119 TTAGCAACAATGCAAGTACGAATCCTCTACAAATCTCCAGCGCAAGATTCAT 2175
Db LeuLeuLeuSerAlaLysGlyGlyAsnAlaGlyAlaProSerAlaGlnValSerSerLeu 910
QY 2176 -----CCTGCAGTCATTTGTAGCACA----- 2196
Db GluAlaLysGlyAsnIleArgLeuValThrGlyLeuThrAspLeuArgGlySerIle 930
QY 2197 ACTGCTGGTCT-----GTTACAATTAGTGGCGCTATCTTTTTCGAGGATTTG 2244
Db ThrAlaGlyLysAsnLeuValValAlaThrThrLysGlyLysLeuAsnIleGluAlaVal 950
QY 2245 GATGATACAGCTTATGATAGGTATGATGCTGCTGCT-----AATCAAAAATC 2295
Db AsnAsnSerPheSerAsnTyPheProThrGlnLysAlaAlaGluLeuAsnGlnLysSer 970
QY 2296 AATGCTCTGAATACAGTTAGGACTAAGCCCAAGCTAATGCCCAATGCTGCTGCTGCT 2352
Db LysGluLeuGlnGlnIleAlaGlnLeuLysLysSerSerProLysSerIleLysLeu 990
QY 2353 -----ACTCTAGGGAATCAGATCCCTAAGTATGGCTATCAAGGAAGCTTGGCTGG 2409
Db ProThrLeuGlnGluArgAspArgLeuAlaPhe----- 1002
QY 2410 GATCCTTAATACGCAATAATGCTCTCTTATCTCTGAAAGCTACATGAGCACTGGG 2469
Db TyrIleGlnAlaIleAsnLysGluValLysGly 1013

QY 2470 TATAATCCT---GGGCTGAGCGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCC 2526
Db LysLysProLysGlyLysGlyLysLeuGlnAla-----LysLeuSerAlaGln 1029
QY 2527 ATTTTATATATACATCTGCGCAT-----TCAGCAATTCAGCAAGT 2568
Db AsnIleAspLeuIleSerAlaGlnGlyIleGluLysSerGlySerAspIleThrAlaSer 1049
QY 2569 GTGATGGGCTCTTATTGTCGAGGA----- 2595
Db LysLysLeuAsnLeuHisAlaAlaGlyValLeuProLysAlaAlaAspSerGluAlaAla 1069
QY 2596 ---TTATGGGTTCTGGAGTTTCGAATTTCTTC----- 2625
Db AlaIleLeuIleAspGlyIleThrAspGlnTyGluIleGlyLysProThrTyLysSer 1089
QY 2626 TATCATGACCGGATCTTTAGGTACGGTATATATAGTGG-----GGTTAT 2679
Db HisTyAspLysAlaAlaLeuAsnLysProSerArg-----LeuThrGlyArgThrGlyVal 1108
QY 2680 TCCTTAGGAGCAAACTCTTACTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAGTA 2739
Db SerIleHisAlaAlaAlaLeuAspAlaArgIleIleGlyAlaSerGluIle 1128
QY 2740 TTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAAATCATCATGCTTGCATAGATCC 2799
Db LysAlaProSer-----GlySer 1134
QY 2800 GTTTATCTATCTACCAACAGCTTTATGTGATCTTCTTCGGAGATCG----- 2853
Db IleAspIleLysAlaHisSerAspIleValLeuGluAlaGlyGlnAsnAspAlaTyThr 1154
QY 2854 TTTATCCGTCTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTGCA 2913
Db PheLeuThrLysGlyLysSerGlyLysIleAlaArgLysThrLysPheThrSerThr 1174
QY 2914 GAGGAG-----AGCATGTTCTGGGATAATACTGCTGCTGCTGAGAG 2958
Db ArgAspHisLeuIleMetProAlaProValGluLeuThrAlaAsnGlyIleThrLeuGln 1194
QY 2959 ATTGGA-----GCGGATTACCGATT 2979
Db AlaGlyGlyAsnIleGluAlaAsnThrThrArgPheAsnAlaProAlaGlyLysValThr 1214
QY 2980 GTGATTACTCCATCTAAGCTCTATTGAATGATGTTGCGCTTCTGTCAGCTGAGTTT 3039
Db LeuValAlaGlyGluGluLeuGlnLeu-----LeuAlaGluGlu 1227
QY 3040 TCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCATTCAGAGC 3099
Db GlyIleHisLysHisGluLeu-----AspValGlnLysSerArgPhe----- 1242
QY 3100 GGACATCTCTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGATCGATGT-----TCT 3153
Db ---GlyIleLysValGlyLysSerAsnTySer 1253
QY 3154 AGTACACATCTCTAATAATATATAGCTTTATGCGCGCTTATATCTGTCATCTTATCGCACC 3213
Db LysAsnGluLeuAsnGluThrLysLeuProValArgValAlaGlnThrAlaAlaThr 1273
QY 3214 ATCTCTGCTGCTGAGCAACGCTCTCTATCCCATCAA----- 3249
Db ArgSerGlyTyPheThrValLeuGluGlyThrGluPheLysThrThrLeuAlaGlyAla 1293
QY 3250 -----GAGCATGACACACAGATGCCCTTCTTATTCAGCAACAT 3288
Db AspIleGlnAlaGlyValGlyGlyLysAlaArgValAspAla---LysIleLeuLys 1312
QY 3289 GAGTTGCTGTTAGGAGTCTATGATGCTTCTTAAACAGTAATATAGAGTATATGCG 3348
Db GlyIleValAsnArgIleGlnSerGluGluLysLeuGluThrAsnSerThrValTrp--- 1331
QY 3349 CATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTTGTAGTGCAGGAGTAGA 3408

```
Db 1332 -----GlnLysGlnAlaGlyArgGlySerThrIleGluThrLeuLysLeu 1346
Qy 3409 GTCCGGTCTCAAAA-----TATTGGTTAGAT 3435
Db 1347 ProSerPheGluSerProThrProProLysLeuSerAlaProGlyGlyIleValAsp 1366
Qy 3436 AGTTAAGTGTAGCGATCGCTTTTCTTTCAGATCTACATCATTTGTTTGTAGCTTGT 3495
Db 1367 -----IleProLysGlyAsnLeuLysThrGluIleGluLysLeuSerLys 1381
Qy 3496 TTGTGTTCTATTGTTAGTATCGCGAGCTCTCTCAAGTGTAAAGCTTAATGTAAACC 3555
Db 1382 GlnProGluIleAlaTyr-----LeuLysGlnLeuGlnValAla 1394
Qy 3556 ACTCTCTTTAAGGAGCAGATGTTTACTTGAATGGAGAC----- 3594
Db 1395 LysAsnIleAsnTrpAsnGlnValGlnLeuAlaTyrAspArgTrpAspTyrLysGlnGlu 1414
Qy 3595 -----TGCGCTTTTGTCAATGCTCTATGCAGGA 3621
Db 1415 GlyLeuThrGluAlaGlyAlaAlaIleIleAlaLeuAlaValThrValThrSerGly 1434
Qy 3622 GCTGAAGAGGTTTCGATTATCTCAGCTAATGCGCAATTTAAAGATTACC----- 3672
Db 1435 AlaGlyThrGlyAlaValLeuGlyLeuAsnGlyAlaAlaAlaAlaThrAspAlaAla 1454
Qy 3673 -----GACAAACCATTACATTATCATTTACAGATTCTCAAGGCCAGTT--- 3717
Db 1455 PheAlaSerLeuAlaSerGlnAlaSerValSerPheIleAsnAsnLysGlyAspValGly 1474
Qy 3718 -----CTTCAAAATTATGCTTCATTTTCAGCA 3744
Db 1475 LysThrLeuLysGluLeuGlyArgSerSerThrValLysAsn-----LeuValValAla 1492
Qy 3745 GGAGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAATATGTTCTTGC 3804
Db 1493 AlaAlaThrAlaGlyValAlaAspLysIleGlyAlaSerAlaLeuAsnAsnValSer--- 1511
Qy 3805 GGAGAAAGGAATGATCTCCGGGAAACCTGAGTATTTCCGGAGCAGGCGAAGTGATT 3864
Db 1512 ---AspLysGlnTrpIleAsnAsnLeuThrValAsnLeuAlaAsnAlaGlySerAlaAla 1530
Qy 3865 TTCTGGGATACTCCGTGGGTATTCTCTTTATCTACTGTGCCAACCTCATCACT 3924
Db 1531 LeuIleAsnThrAlaIleAsnGlyGlySerLeuLys----- 1542
Qy 3925 CCGCTGCTCCAACAGTTAGTATGCTCGGAAAGGCTATTTTTTCTGTAGAGACTAGT 3984
Db 1543 -----AspAsnLeuGlyAspAlaAlaLeuGlyAlaIleValSerThrValHisGly 1559
Qy 3985 TTGGAGATCTCAGGCGTCAAAAAGGGTCTGTTTCGATAATAATGCCGGGAATTTCCGA 4044
Db 1560 GluValAlaSerLysIleLysPheAsnLeuSerGluAspTyrIleThrHisLysIleAla 1579
Qy 4045 ACAGTTTTCAGGT-----AAGAATAATAATAATGCTGGTGGTGGAGGC 4089
Db 1580 HisAlaIleAlaGlyCysAlaAlaAlaAlaAlaAlaLysGlyLysCysGlnAspGly-Al 1599
Qy 4090 AGTGGGTTCCGCTACA-----CCATCAAGTAC 4116
Db 1599 aIleGlyAlaAlaValGlyGluIleValGlyGluAlaLeuThrAsnGlyLysAsnProAl 1619
Qy 4117 GACTTTTACAGTTAAACCTGTAAGGGAAGTTCTTTCACAGATAACGTAGCTCTTG 4176
Db 1619 aThrLeuThrAlaLysGluArgGluGlnIleLeuAlaTyrSerLysLeuValAlaGlyTh 1639
Qy 4177 CGAGCGCGAGTGGTT 4192
Db 1639 rValSerGlyValVal 1644
```

RESULT 35

US-11-045-208-38

```
; Sequence 38, Application US/11045208
; Publication No. US20060063167A1
; GENERAL INFORMATION:
; APPLICANT: (A) NAME : I.N.S.E.R.M
; TITLE OF INVENTION: DNA, specific proteins and peptides
; of the Neisseria meningitidis species bacteria, method
; for obtaining them and their biological application.
;
; NUMBER OF SEQUENCES: 99
; STREET: New Horizons Court
; CITY: Brentford
; COUNTRY: Great-Britain
; ZIP: TW 89 EP
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OBB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/045,208
; FILING DATE: 31-Jan-2005
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/214,759
; FILING DATE: 22-APR-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: SMITHKLINE BEECHAM
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1981 acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..1981
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-11-045-208-38

Alignment Scores:
Pred. No.: 1,73e-08 Length: 1981
Score: 199.00 Matches: 306
Percent Similarity: 32.5% Conservative: 200
Best Local Similarity: 19.6% Mismatches: 591
Query Match: 2.5% Indels: 461
DB: 71

US-10-701-844-1 (1-4435) x US-11-045-208-38 (1-1981)
Qy 448 GGGGGGGATATGACGCGAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTA 507
Db 284 GlyGlyMetTyrAlaAspSerIle----- 291
Qy 508 ACTGTATCATTTCCCTATACCTACTGTTATAGAGATCCGAGTGGGACTACTGTTTCTGCA 567
Db 292 -----ThrLeuIleAlaAsnGlnLysGlyValGlyValLysAsnAla 305
Qy 568 GGAGAGTTAACAATAAAAAATCTTGACAATTTCTATTCGACGCTTTGCTTAAAGTTGTTT 627
Db 306 GlyThrLeu-----GluAlaAlaLysGlnLeuIleValThrSerSer 319
Qy 628 GGGAACTATTATGGGAGT-----TTTACTGTTTGGGAGAGACACTCGTTG 675
Db 320 GlyArgIleGluAsnSerGlyArgIleAlaThrThrAlaAspGlyThrGluAlaSerPro 339
Qy 676 ACTTTCGAGAACATACGAGCTTCTACAAATGGGCGAGCT-----CTAAGTAATAGC 726
Db 340 ThrThrLeuSerIleGluThrThrGlnLysGlyAlaAlaGlyThrPheIleSerAsnGly 359
Qy 727 GCT-----GCTGATGGACTGTTTACTATTAGAGGTTTAAAGAAATATCTTTTCC 777
Db 360 GlyArgIleGluSerLysGlyLeuValIleGluThrGlyGluAspIleSerLeuArg 379
Qy 778 AATGCAATTTCACTTCTGCGCTACTGCTGCTGCAACGACTAATAAGGATGACCGACT 837
```

```
Db 380 Aen-----GlyAlaValValGluAsnAsnGlySerArgPro 391
QY 838 CGAGCACAACA---TCTACACCGTCTAATAGTACTATTATTCTTAAACACAGATCTTTTG 894
Db 392 AlaThrThrValLeuAsnAlaGlyHisAsnLeuValIleGluSerLysThrAsn----- 409
QY 895 TTACTCAATAATGAGAGTTCTCATTCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCT 954
Db 410 ---ValAsnAsnAlaLysGlySer-----AlaAsnLeuSerAla-----GlyGlyArg 424
QY 955 ATAGATGCTAAGAGCTTAAAGGTTCAAGCAAGTTAGCAAGCTTTGTCTCTTCCAGCAAAAT 1014
Db 425 ThrThrIleAsnAspAlaThrIleGlnAlaGlySerSerValTyR-SerSerThrLysGly 444
QY 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTCTATGGCTAAC 1074
Db 445 AspThrGluLeuGlyGluAsnThrArgIleIleAlaGluAsnValThrValLeuSerAsn 464
QY 1075 GAGGCTCCTATTGCTTGTAGCGAATGTTGCGAGGAGTAGAGGGGAGGATTCGTGCT 1134
Db 465 GlySer-----IleGlySerAlaAlaValIleGluAlaLysAspThrAlaHis 480
QY 1135 GTTCAGGATGGGCGAG-----CAGGGAGTGTCACTCATCTTCAACA 1176
Db 481 IleGluSerGlyLysProLeuSerLeuGluThrSerThrValAlaSerAsnIleArgLeu 500
QY 1177 GAAGATCCAGTAGTAAGTTTTCAGAAATACCTCGGTAGAGTTTGTATGGGACGTAGCC 1236
Db 501 AsnAsnGlyAsnIleLysGlyLysGlnLeuAlaLeuAlaLeuAlaAspAsnIleThr 520
QY 1237 CGAGTAGGAGGAGGATTTACTCTCGGAAAGTTGCTTCTCTGTAATATGAGAAAC 1296
Db 521 AlaLysThrThrAsnLeuAsnThrProGlyAsnLeu---TyrValHisThrGlyLysAsp 539
QY 1297 TTGTTTCTCAAC-----AATGTTGCTCTCTCTGTTTACATTCGTCTGCAACCAACA 1350
Db 540 LeuAsnLeuAsnValAspLysAspLeuSerAlaAlaSerIleHisLeuLysSerAspAsn 559
QY 1351 AGTGGACAGGCTTCTAATACAGATAATAATTACGGAGATGAGGAGCTATCTCTGTAAG 1410
Db 560 AlaAlaHisIleThrGlyThrSerLysThrLeuThrAlaSerLysAspMet----- 576
QY 1411 AATGTCGCGCAAGCAGA-----TCCAATACTCTGGA 1443
Db 577 ---GlyValGluAlaGlyLeuLeuAsnValThrAsnThrAsnLeuArgThrAsnSerGly 595
QY 1444 TCAGTTCTCTTGTGAGGAGGAGTAGTTTCTTTAGTACCAATGATGCTGTGGAAA 1503
Db 596 AsnLeu-----HisIleGlnAlaAlaLys 603
QY 1504 GGGGGA---GCTATTATTATGCAAAAAGCTCTCGTTTCTAACTGTGTCCTGTACAAATTT 1560
Db 604 GlyAsnIleGlnLeuArgAsnThrLysLeuAsnAlaAlaLysAlaLeuGluThrThrAla 623
QY 1561 TTAAGG---AATATCGCTAATATGATGGGCGGATTTATTATTAGAGAAATCTGGAGAGCTC 1617
Db 624 LeuGlnGlyAsnIleValSerAsp---GlyLeuHisAlaValSerAlaAspGlyHisVal 642
QY 1618 AGTTTATCTGCTGATTATGAGATATTATTTCGATCGGAATCTTAAAGACAGCCAAA 1677
Db 643 SerLeuLeuAlaAsn---GlyAsnAlaAspPheThrGlyHisAsnThrLeuThrAlaLys 661
QY 1678 GAGAAATGCTCGCATGTTAATGGC----- 1701
Db 662 -----AlaAspValAsnAlaGlySerValGlyLysGlyArgLeuLysAlaAspAsn 678
QY 1702 ---GTAATCTGTCTCTCACAAGCCATTTTCGATG 1731
Db 679 ThrAsnIleThrSerSerGlyAspIleThrLeuValAlaGlyAsnGlyIleGlnLeu 698
QY 1732 GGATCGGGA-----GGGAAATACGACATTAAGAGCTAAAGCA 1770
```

```
699 GlyAspGlyLysGlnArgAsnSerIleAsnGlyLysHisIleSerIleLysAsnAsnGly 718
1771 GGGCATCAGATTTCTTTAATGATCCCATCGAGATGGCAACCGA----- 1815
Db 719 GlyAsnAlaAspLeuLysAsnLeuAsnValHisAlaLysSerGlyAlaLeuAsnIleHis 738
1816 -----AATAACCGCAGCGCAGCTCTCC----- 1839
Db 739 SerAspArgAlaLeuSerIleGluAsnThrLysLeuGluSerThrHisAsnThrHisLeu 758
1840 -----AAACTTCTAAAAATTAAACGATGCTGAAGGATAC----- 1872
Db 759 AsnAlaGlnHisGluArgValThrLeuAsnGlnValAspAlaTyAlaHisArgHisLeu 778
1873 -----ACAGGGGATATTGTTTT----- 1890
Db 779 SerIleThrGlySerGlnIleTrpGlnAsnAspLysLeuProSerAlaAsnLysLeuVal 798
1891 GCTAATGGA-----ACAGTACTTTTGTATCCAAAATGTTACGATA 1929
Db 799 AlaAsnGlyValLeuAlaLeuAsnAlaArgTyrSerGlnIleAlaAspAsnThrThrLeu 818
1930 GAGCAAGGAGGATTTCTTCGTGAAAGGCAAAATTA-----TCAGTGAAT 1977
819 ArgAlaGlyAlaIleAsnLeuThrAlaGlyThrAlaLeuValLysArgGlyAsnIleAsn 838
1978 -----TCTTAAGTCAGACAGGTGG 1998
839 TrpSerThrValSerThrLysThrLeuGluAspAsnAlaGluLeuLysProLeuAlaGly 858
1999 AGTCTGTATATGAAAGCTGGAGTACATGGGATTTTGTAACTCCACCAACCCACCAACAG 2058
859 ArgLeuAsnIleGluAlaGlySer-----GlyThrLeuThrIleGluPro----- 873
2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCATCTGCATTTGCTCTTCTTCTTTG 2118
874 -----AlaAsnArgIleSerAlaHisThrAspLeuSerIleLysThrGlyLys 890
2119 TTAGCAACAATGCGAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTTCTCAT--- 2175
891 LeuLeuLeuSerAlaLysGlyGlyAsnAlaGlyAlaProSerAlaGlnValSerSerLeu 910
2176 -----CCTGCAGTCATTGGTAGCACA----- 2196
911 GluAlaLysGlyAsnIleArgLeuValThrGlyGluThrAspLeuArgGlySerLysIle 930
2197 ACTGCTGGTCT-----GTTACAATTAGTGGCGCTATCTTTTGGAGATTG 2244
931 ThrAlaGlyLysAsnLeuValValAlaThrThrLysGlyLysLeuAsnIleGluAlaVal 950
2245 GATGATACAGCTTATCATAGTATGATTGCTAGGTTCT-----AATCAAAAATC 2295
951 AsnAsnSerPheSerAsnTyrPheProThrGlnLysAlaAlaGluLeuAsnGlnLysSer 970
2296 AATGCTCTGAAATACAGATTTAGGAGCTAAAGCCCCAGCTAATGCCCATCAGATTG--- 2352
971 LysGluLeuGluGlnGlnIleAlaGlnLeuLysLysSerSerProLysSerLysLeuIle 990
2353 ---ACTCTAGGGAATCAGATCGCTTAAGTATGCTATCAAGAAAGCTGGAAAGCTTGGCTGG 2409
991 ProThrLeuGlnGluArgAspArgLeuAlaPhe----- 1002
2410 GATCCTAATACCAAAATAATGTCTCTTATCTCTGAAAGCTACATGAGCTAAAGCTGGG 2469
1003 -----TyrIleGlnAlaIleAsnLysGluValLysGly 1013
2470 TATAATCT---GGGCTGAGCGAGTAGTCTTTTGGTTCCAAATAGTTTATGGGATCC 2526
1014 LysLysProLysGlyLysGluTyrLeuGlnAla-----LysLeuSerAlaGln 1029
2527 ATTTTATAGATACATCTGCGCAT-----TCAGCAATTCAGCAAGT 2568
1030 AsnIleAspLeuIleSerAlaGlnGlyIleGluIleSerGlySerAspIleThrAlaSer 1049
```


QY 2569 GTGGATGGCGCTCTATTATGCGAGG----- 2595
Db LysLysLeuAsnLeuHisAlaAlaGlyValLeuProLysAlaAlaAspSerGluAlaAla 1069
QY 2596 ---TTATGGGTTTCGGAGTTTCGAATTCCTC----- 2625
Db AlaIleLeuIleAspGlyIleThrAspGlnIleGlyLysProThrTyrLysSer 1089
QY 2626 TATCATGACCGCATGCTTTAGGTACGGATATCGTATATAGTGGG-----GGTTAT 2679
Db HisTyrAspLysAlaAlaLeuAsnLysProSerArg---LeuThrGlyArgThrGlyVal 1108
QY 2680 TCCTTAGAGGAACTCCTACTTTGGATCATCGATGTTTGGTCTAGCATTTACCAAGTA 2739
Db SerIleHisAlaAlaAlaLeuAspAlaArgIleIleGlyAlaSerGluIle 1128
QY 2740 TTTGGTAGATCTAAAGATTATGTAGTGTGCTGCTTCCAATCATCATGCTTGATAGATCC 2799
Db LysAlaProSer-----GlySer 1134
QY 2800 GTTTATCTATCTACCAACAGCTTTATGTGATCTATTGTCGGAGATGCG----- 2853
Db IleAspIleLysAlaHisSerAspIleValLeuGluAlaGlyGlnAsnAspAlaTyrThr 1154
QY 2854 TTTATCCGTGCTAGCTACGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCA 2913
Db PheLeuLysThrLysGlyLysSerGlyLysIleAlaArgLysThrLysPheThrSerThr 1174
QY 2914 GAGGAG-----AGCATGTTCTGTTGGATATAAATGCTGTGCTGGCTGAGAG 2958
Db ArgAspHisLeuIleMetProAlaProValGluLeuThrAlaAsnGlyIleThrLeuGln 1194
QY 2959 ATTGA-----CGCGGATTACCGATT 2979
Db AlaGlyGlyAsnIleGluAlaAsnThrThrArgPheAsnAlaProAlaGlyLysValThr 1214
QY 2980 GTGATTACTCCATCTAAGCTCTATTGTAATGATGGTCTCTTTCGTGCAAGCTGAGTTT 3039
Db LeuValAlaGlyGluGluLeuGlnLeu-----LeuAlaGluGlu 1227
QY 3040 TCTTATGCCGATCATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCATTCAGAGC 3099
Db GlyIleHisLysHisGluLeu-----AspValGlnLysSerArgPhe----- 1242
QY 3100 GGACATCTCTAAATCTATCACTCTCTGTTGGAGTCAAGTTTGTATCGATGT-----TCT 3153
Db -----IleGlyIleLysValGlyLysSerAsnTyrSer 1253
QY 3154 AGTACACATCTTAATAATATAGCTTTATGCGGGCTTATATCTGTAGTCTTATCGCAC 3213
Db LysAsnGluLeuAsnGluThrLysLeuProValArgValValAlaGlnThrAlaAlaThr 1273
QY 3214 ATCTCTGTTACTGAGCAAGCTCTCTATCCATCA----- 3249
Db ArgSerGlyTyrAspThrValLeuGluGlyThrGluPheLysThrThrLeuAlaGlyAla 1293
QY 3250 -----GAGACATGCACACAGATGCTCTTCATTTAGCAAGACAT 3288
Db AspIleGlnAlaGlyValGlyGluLysAlaArgValAspAla---LysIleIleLysLys 1312
QY 3289 GGAGTTGTGGTTAGAGATCTATGATGCTTCTTAACAAGTAATATAGATATATGCG 3348
Db GlyIleValAsnArgIleGlnSerGluGluLysLeuGluThrAsnSerThrValTyr--- 1331
QY 3349 CATGGAAGATATGAGTATCGAGATGCTTCTCGAGCTATGTTTGTAGTCAGAGATAGA 3408
Db -----GlnLysGlnAlaGlyArgGlySerThrIleGluThrLeuLysLeu 1346
QY 3409 GTCCGGTCTAAAA-----TATTGGTTAGAT 3435
Db ProSerPheGluSerProThrProProLysLeuSerAlaProGlyGlyTyrIleValAsp 1366

QY 3436 AGTTAAGTGTAGCGATGCTCTTTTCTTTGAGATCTACATCATTTTGTCTTTTACCTGT 3495
Db -----IleProLysGlyAsnLeuLysThrGluIleGluLysLeuSerLys 1381
QY 3496 TTGTGTTCTTATCGTATGATTCGCGAGCTCTCTCAAGTGTAAACCCCTAATAACCC 3555
Db GlnProGluTyrAlaTyr-----LeuLysGlnLeuGlnValAla 1394
QY 3556 ACTCCTTTTAAAGGAGAGCATGTTTACTTTGAATGGAGAC----- 3594
Db LysAsnIleAsnTyrAspArgTyrAspArgTyrAspTyrLysGlnGlu 1414
QY 3595 -----TGCGCTTTTGTCAATGTCTATCGAGA 3621
Db GlyLeuThrGluAlaGlyAlaAlaIleAlaLeuAlaValThrValValThrSerGly 1434
QY 3622 GCTGAAGAGGTTTCGATTATCTCACTCAATCGACCAATTTTAAACGATTACC----- 3672
Db AlaGlyThrGlyAlaValLeuGlyLeuAsnGlyAlaAlaAlaAlaThrAspAlaAla 1454
QY 3673 -----GGCAAAACCATATCATTTACAGATCTTCAAGGGCCAGTTCTT 3720
Db PheAlaSerLeuAlaSerGlnAlaSerValSerPheIleAsnAsnLysGlyAspVal--- 1473
QY 3721 CAAATTTATGCTTCATTTTCAGCAGGAGACACTTACTCTGAGAGATTTTTCGAGTCTG 3780
Db -----GlyLysThrLeuLys----- 1478
QY 3781 ATGTTCTCGAAAAATGTTTCTTGGCGAAGGAATGATCTCCGGGAAACCCGTGAGT 3840
Db -----GluLeuGlyArgSerThrValLysAsnLeuValValAla 1492
QY 3841 ATTTCCGAGCAGCGAAGTGAATTTTCGGATAACTCCGGGGTATTTCTCTTTATCT 3900
Db AlaAlaThrAlaGlyValAla-----AspLysIleGlyAlaSerAlaLeuAsn 1508
QY 3901 ACTGTGCCAATCATCATCACTCCGCTGCTCCACAGATTAGTGTGCGGAAAGGG 3960
Db AsnVal---SerAspLysGlnTyrIleAsnAsnLeuThrValAsnLeuAlaAsnAlaGly 1527
QY 3961 TCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAAAAAGGGTCTATGTT 4020
Db SerAlaAlaLeuIleAsnThrAlaVal----- 1536
QY 4021 GATAATAATGCGGAATTTCCGAACAGTCTTTTCGAGTAAGATAATAATAATCTGCT 4080
Db -----AsnGlyGlySerLeu-LysAspAsnLeuGluAlaAsnIleLeuAlaAlaLeuVa 1554
QY 4081 GGTGAGGAGCAGTGGTTCGCTTACACCATCAAGTACGACTTTTACAGTTAAAACTGTAA 4140
Db AsnThrAlaHisGlyGluAlaAlaSerLysIleLysGlnLeuAspGlnHisTyrIleVa 1574
QY 4141 AGGGAAGTCTTTTTCACAGATAAGTACCTTTCGCGAGCGGAGTGGTTTATAAAGG 4200
Db HisLysIleAlaHisAla-----IleAlaGlyCysAlaAlaAlaAlaAsnLysGly 1592
QY 4201 CATCTGTCTTTTCAAGACAATGAAGGAGGCGATA----- 4234
Db -----LysCysGlnAspGlyAlaIleGlyAlaAlaValGlyGluIleValGlu 1608
QY 4235 -----TCTTCCGAGGGAAC-----ACGATACGATGATTTAAGGAT 4272
Db yGluAlaLeuThrAsnGlyLysAsnProAspThrLeuThrAlaLysGluArgGluGlnIle 1628
QY 4273 TCTCTCTGCTACTACAGTACAGATACAGATACAGAGAGCGGCGGTGGAGGA 4324
Db eLeuAlaTyrSerLysLeuValAlaGlyThrValSerGlyValValGlyGly 1645

RESULT 36

US-11-067-260-20

; Sequence 20, Application US/11067260

; Publication No. US20060051840A1

; GENERAL INFORMATION:


```

Db      783  LeuArgValIleAlaGlnThrGlnAspGlyGlyThrTrpGluGlnGlyGlyValGlu 802
QY      2032  -----TTTGAACCTCCACCAACCAACAGCCT 2061
Db      803  GlyLysMetArgGlySerThrGlnThrValGlyIleAlaAlaLysThrGlyGluAsnThr 822
QY      2062  CCTGGCGCTAATCAAGTTCAGCTTCAGCTTCCCAATCGCATTTGTCTCTTCTTCTTGTGA 2121
Db      823  ThrAlaAla-----AlaThrLeuGlyMetGlyArgSerThrTrpSer 836
QY      2122  GCAACAATGCAATGACGATTACGAATCTCTACCAATCTCCAGCGCAAGATTCCTCATCTGCA 2181
Db      837  GluAsnSerAla-----AsnAlaLysThr-----AspSerIleSerLeu 849
QY      2182  GTCATTGGT---AGCACAACTGCTGGTCTCTGTACAAATAGTAGGGCCCTATCTTTTTCAG 2238
Db      850  PheAlaGlyIleArgHisAspAlaGlyAspIleGlyTyLeuLysGlyLeuPhe----- 867
QY      2239  GATTGGATGATACAGCTTATGATAGGTATGAT-----TGGCTAGGTCT 2283
Db      868  -----SerTyrglyArgTyrglyAsnSerIleSerArgSerThrGlyAla 882
QY      2284  AATCAA-----AAATCAATGTCCTGAAATACAGTTAGGGAAGTAAGCCCCA 2331
Db      883  AspGluHisAlaGluGlySerValAsnGlyThrLeuMetGlnLeuGlyAlaLeuGlyGly 902
QY      2332  GCTAATGCCCA-----TCAGATTGACTCTAGGAAAT----- 2364
Db      903  ValAsnValProPheAlaAlaThrGlyAspLeuThrValGluGlyLeuArgTyrglyAsp 922
QY      2365  -----GAGATCCCTAAGTATGCTATCAAGAACTGGAAGCTTGGTGGATCCTAAT 2418
Db      923  LeuLeuLysGlnAspAlaPheAlaGluLysGlySer---AlaLeuGlyTyrglyAsn 941
QY      2419  ACAGCAATATGTCCTTATCTCTGAAAGCTACATGAGCTAAAGCTGGGTATATCT 2478
Db      942  SerLeuThrGluGly----- 946
QY      2479  GGGCCTGAGCGAGTAGCTTCTTGGTCCAAATAGTTATGGGATCCATTTTAGAT--- 2535
Db      947  -----ThrLeuValGlyLeuAlaGlyLeuLysLeuSerGlnProLeuSerAspLys 963
QY      2536  ---ATACGATCTGGCATTCAGCAATTCAGCAAGTGTGGATGGGGCTCTTATTGTGCA 2592
Db      964  AlaValLeuPheAlaThrAlaGlyValGluArgAspLeuAsnGlyArgAspTyrglyThr--- 982
QY      2593  GGATTATGGTTCTGGAGTTTCGAATTTCTCTATCATGACCGCGATGCTTTAGTCAG 2652
Db      983  -----ValThrGlyGlyPheThrGlyAlaThrAlaAlaThrGlyLys 996
QY      2653  -----GGATATCGGTATATTTAGTGGGGTTATTCCTTAGGAGCA 2691
Db      997  ThrGlyAlaArgAsnMetProHisThrArgLeuValAlaGly-----LeuGlyAla 1013
QY      2692  AACTCCTACTTTGGATCATCGATTTGTGCTAGCATTTACCGAAGTATTGGTAGATCT 2751
Db      1014  AspValGluPheGlyAsnGlyTrpAsnGlyLeuAlaArgTyrglySerTyrglyAlaGly---Ser 1032
QY      2752  AAAGATTATGATGTGTGCTTCCCAATCATCATGCTTGCATAGGATCCGTTTATCTATCT 2811
Db      1033  LysGlnTyrgly-----GlyAsnHisSerGlyArgValGly----- 1043
QY      2812  ACCCAACAAGCTTTATGTGGATCCTATTTGTCGGAGATGCGTTTATCGTGTAGCTAC 2871
Db      1044  -----ValGlyTyrglyArgPheLeuGluGlySerGly 1053
QY      2872  GGGTTTGGAAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCATGTTCT 2931
Db      1054  GlyGlyGly----- 1056
QY      2932  TGGGATAATAACTGTCTGCTGGTGGAGAGATTGGAGCGGATTTACCGATTGTGATTACTCCA 2991

```

RESULT 37

US-11-067-260-18

; Sequence 18, Application US/11067260

```

Db      1057  -----ValAlaAlaAspIleGlyAlaGlyLeuAlaAspAlaLeuThrAla 1071
QY      2992  TCTAAGCTCTATTGAATGAGTTCGCTTCTTCGCAAGCTGAGTTTCTTATGCCGAT 3051
Db      1072  ProLeuAspHisLysAspLysGlyLeuGlnSerLeuThrLeuAspGlnSerValArgLys 1091
QY      3052  CATCAATCTTTTACAGAGGAGCGCATCAAGCT---CGGGCATTCACAGAGCGGATCTC 3108
Db      1092  AsnGluLysLeuLysLeuAlaAlaGlnGlyAlaGluLysThrTyrglyAsnGlyAspSer 1111
QY      3109  CTAATCTATCAGTTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACATCCTAAT 3168
Db      1112  LeuAsnThrGly-----LysLeuLysAsnAspLysValSer----- 1123
QY      3169  AAATATAGCTTTATGGCGGCTTATATCTGTGATGCT----- 3204
Db      1124  ArgPheAspPheIleArgGlnIleGluValAspGlyGlnLeuIleThrLeuGluSerGly 1143
QY      3205  ---TATCCACCATCTCTGGTACTGAGACAACGCTCCTATCCCATCAAGAGACATGGACA 3261
Db      1144  GluPheGlnValTyrglyLysGlnSerHisSerAlaLeuThrAlaPheGlnThrGluGlnIle 1163
QY      3262  ACAGATGCTTTCATTTAGCAAGACATGGAGTGTGTGTTAGA-----GGA 3306
Db      1164  GlnAspSerGluHisSerGlyLys-----MetValAlaLysArgGlnPheArgIleGly 1181
QY      3307  TCTATGATGCTTCTTAAACAAGTAATATAGAGTATATGCGCATGGAAGATATGATGAT 3366
Db      1182  AspIleAlaGlyGluHisThrSerPheAspLysLeuProGluGlyGlyArgAlaThrTyrgly 1201
QY      3367  CGAGATGCTTCTCGAGGCTATGGTTTGAGTCAGCAAGTAGAGTCCGGTCTTAAAAATAT 3426
Db      1202  ArgGlyThrAlaPheGlySerAsp---AspAlaGlyGlyLysLeu----- 1215
QY      3427  TGGTTAGATAGTTAAGTGTAGCGATGCTTTTCTTTGAGATCTACATCATTTTGTTTT 3486
Db      1215  ----- 1215
QY      3487  TTAGCTTGTGTGTTCCTATTCGTATGATTCGGAGC----- 3525
Db      1216  -----ThrTyrglyThrIleAspPheAlaAlaLysGlnGlyAsnGlyLysIle 1230
QY      3526  -----TCTCCTCAAGTGTTAACGCGCTAATGTAACCACTCCTCTTTTAAGGAGAC 3573
Db      1231  GluHisLeuLysSerProGlu----- 1237
QY      3574  GATGTTTACTTGAATGGAGACTGCGCTTTTGTCAATGTCTATGCGAGGAGCTGGAAGGT 3633
Db      1238  -----LeuAsnValAspLeuAlaAlaAlaPheIleLysProAspGlyLysArgHis 1254
QY      3634  TCGATTATCTCAGCTAATGGGACAAATTAACGATTACGGGACAAAACCATACATTATCA 3693
Db      1255  AlaValIleSer----- 1258
QY      3694  TTTACAGATTCTCAAGGGCCAGTTCCTCAAAATATGCTTCATTTTCAGCAGGAGAGACA 3753
Db      1259  -----GlySerValLeuTyrglyAsnGlnAla----- 1266
QY      3754  CTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCTTTCGCGAGAAAG 3813
Db      1267  -----GluLys 1268
QY      3814  GGAATGATCTCC-----GGGAAAAACGGTAGTATTTCCGAGCAGCGGAA 3858
Db      1269  GlySerTyrglySerLeuGlyIlePheGlyGlyLysAlaGlnGluValAlaGlySerAlaGlu 1288
QY      3859  GTG 3861
Db      1289  Val 1289

```

; Publication No. US20060051840A1			
; GENERAL INFORMATION:			
; APPLICANT: Arico, Maria			
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISSERIAL PROTEINS			
; FILE REFERENCE: 22300209501			
; CURRENT APPLICATION NUMBER: US/11/067,260			
; CURRENT FILING DATE: 2005-02-25			
; PRIOR APPLICATION NUMBER: 10/220,480			
; PRIOR FILING DATE: 2002-08-28			
; PRIOR APPLICATION NUMBER: PCT/IB01/00420			
; PRIOR FILING DATE: 2001-02-28			
; NUMBER OF SEQ ID NOS: 121			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 18			
; LENGTH: 1474			
; TYPE: PRT			
; ORGANISM: Artificial			
; FEATURE:			
; OTHER INFORMATION: deletion G983-ORF46.1			
US-11-067-260-18			
Alignment Scores:			
Pred. No.:	4,13e-08	Length:	1474
Score:	194.00	Matches:	234
Percent Similarity:	32.1%	Conservative:	139
Best Local Similarity:	20.1%	Mismatches:	438
Query Match:	2.5%	Indels:	352
DB:	11	Gaps:	59
US-10-701-844-1 (1-4435) x US-11-067-260-18 (1-1474)			
QY	559	TTTTCTGCAGGAGTTACATTAAAAATCTTGACATCTTATTCAGCTTTGCGCTTTA	618
DB	268	TyrSerGlycylAspLysThr-----AspGluGlyIleArgLeuMetGlnGln	283
QY	619	AGTTGTTTGGGAACCTATTAGGGAGTTTACTGTTTGTAGGAGAGACACTCGTTGACT	678
DB	284	SerAspTyrGlyAsnLeu-----SerTyrHisIleArgAsnLysAsnMetLeuPheIle	301
QY	679	TTC-----GAGAACATACCGACTTCTACAAATGGGCGACTCTA-----	717
DB	302	PheSerThrGlyAsnAlaGlnAlaGlnProAsnThrTyrAlaLeuLeuProPheTyr	321
QY	718	AGTAATAGCCCTCCTGATGAGCTGTTTACTTATTGAGGCT-----	756
DB	322	GluLysAspAlaGlnLysGlyIleIleThrValAlaGlyValAspArgSerGlyGluLys	341
QY	757	TTTAAAGAAATATCTTT-----TCCAAATTCG	783
DB	342	PheLysArgGluMetTyrGlyGluProGlyThrGluProLeuGluTyrGlySerAsnHis	361
QY	784	AATTCTATTCTCCGCTACTGCTGCTGCTCAACGACTAATAAGGTAGCCAGACTCCGACG	843
DB	362	CysGlyIleThrAlaMetTyrCysLeuSerAlaProTyrGluAlaSerValArgPheThr	381
QY	844	ACAACATCT-----ACACCGTCTAATCGTACTATTATTATTCTAAACA	885
DB	382	ArgThrAsnProIleGlnIleAlaGlyThrSerPheSerAlaProIleValThrGlyThr	401
QY	886	GATCTTTTGTACTCAATATAGAACTCTCATCTATAGTAATTTAGTCTCTGGAGAT	945
DB	402	AlaAlaLeuLeuLeu-----GlnLysTyrProTyrMetSerAsn-----	414
QY	946	GGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTC	1005
DB	415	-----AspAsnLeuArgThrThr	420
QY	1006	CAAGAAATACTGCTCAAGCTGATGAGGGGAGCTTGTCAAGTAGTCACAGTTTC-----	1059
DB	421	LeuLeuThrThrAlaGln-----AspIleGlyAlaValGlyValAspSerLysPheGlyTyr	439
QY	1060	-----TCTGCTATGGCTAAACGAGGCTCCTATTGCTCTTTGTAGCGAAT	1101

Db	440	GlyLeuLeuAspAlaGlyLysAlaMetAsnGlyProAlaSerPheProPheGlyAspPhe	459
QY	1102	GTTGCAGGAGTAAGAGGGGAGGATTGCTGCTGTTCCAGGATGGCAGCAGGGAGTGTC	1161
Db	460	ThrAlaAspThrLysGly-----	465
QY	1162	TCATCTACTTCAACAGAAAGATCCAGTAGTAGTTTTCAGAAATACATCGCGTAGAGTTT	1221
Db	466	-----ThrSerAspIleAlaTyrSerPhe---ArgAsnAspIleSerGlyThr	480
QY	1222	GATGGAACTAGCCCGAGTAGGAGGAGGATTACTCTACGGAACGCTGCTCTTC---	1278
Db	481	GlyGlyLeuIleLysLysGlySerGlnLeuGlnLeuHisGlyAsnAsnThrTyrThr	500
QY	1279	-----CTGAATAATGGAACCTTTGTTTCTCAACAAATGTTGCTTCTCTCTGTT	1326
Db	501	GlyLysThrIleIleGluGlySerLeuValLeuTyrGlyAsnAsnLysSerAspMet	520
QY	1327	TACATTGCTGTAGCAACCAACA-----AGTGACAGGCTTCTAATACGAGTAATAAT	1380
Db	521	ArgValGluThrLysGlyAlaLeuIleTyrAsnGlyAlaAlaSerGlyGlySerLeuAsn	540
QY	1381	TACGAGATGGAGGAGCTATCTTCTGTAAGATGTGCGCAGCAGGATCCAACTACT	1440
Db	541	-----SerAspGlyIleValTyrLeuAlaAspThrAspGlnSerGlyAlaAsnGluThr	558
QY	1441	-----GGATCAGTTTCTTGTATGAGAGGAGGAGTAGTTTCTTTAGTAGCAAT	1488
Db	559	ValHisIleLysGlySerLeuGlnLeuAspGlyLysGlyThrLeuTyr-----	574
QY	1489	GTAGCTGCTGGAAA-----GGGGGAGCTATTTAT	1518
Db	575	ThrArgLeuGlyLysLeuLysValAspGlyThrAlaIleIleGlyGlyLysLeuTyr	594
QY	1519	GCCAAAAG-----	1527
Db	595	MetSerAlaArgGlyLysGlyAlaGlyTyrLeuAsnSerThrGlyArgArgValProphe	614
QY	1528	CTCTGGTTGCTAACTGTGCG---CCTGTACAAATTTTAAAGGAATATCGTAATGATGGT	1584
Db	615	LeuSerAlaAlaLysIleGlyGlnAspTyrSerPhePheThrAsnIleGluThrAspGly	634
QY	1585	GGAGCGATTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATATGAGAT---	1641
Db	635	GlyLeuLeuAlaSerLeuAspSerValGluLysThrAlaGlySerGlu---GlyAspThr	653
QY	1642	-----ATTATTTTCATGGGAATCTTAAAGAACAGCCAAA-----	1677
Db	654	LeuSerTyrTyrValArgArgGlyAsnAlaAlaArgThrAlaSerAlaAlaAlaHisSer	673
QY	1678	-----GAGAACTGCTGCGGATGTTAATGGCTAACTGTGCTCTCACAAGCC	1722
Db	674	AlaProAlaGlyLeuLysHisAlaValGluGlnGlySerAsnLeuGluAsnLeuMet	693
QY	1723	ATTTCGATGGGA-----TCGGGAGGGGAAATAACGACATTAAGAGCTAAAGCAGGG	1773
Db	694	ValGluLeuAspAlaSerGluSerSerAlaThrProGluThrValGluThrAlaAla	713
QY	1774	CATCAGATTCTCTTAAATGATCCATCGAGATGGCAACCGAAATACCGCAGCGCAG	1833
Db	714	-----AspArgThrAspMetProGlyIleArgProTyrGlyAlaThr	727
QY	1834	TCTTCCAACTTCTAAAATTACGATGGTGAAGGATACACAGGGGATATGTTTGTGCT	1893
Db	728	PheArgAlaAlaAlaValGlnHisAlaAsnAlaAlaAspGlyValArgIlePheAsn	747
QY	1894	AATGGAGCAGTAGTACTTTGTACCAAAATGTTTACCATAGAG-----CAAGGAAGG	1941
Db	748	SerLeuAlaAlaThrValTyrAlaAspSerThrAlaAlaHisAlaAspMetGlnGlyArg	767
QY	1942	ATTGTTTCTCGTGAAGGCAAAATATATAGTGAATTTCTCTAAGTCAGACAGGTGGAGT	2001
Db	768	-----ArgLeuLysAlaValSerAspGlyLeuAspHisAsnGlyThrGly	782

```
QY 2002 CTGTATATGGAAGCT-----GGGAGTACATGGAT----- 2031
Db 783 LeuArgValIleAlaGlnThrGlnGlnAspGlyGlyThrTrpGluGlnGlyValGlu 802
QY 2032 -----TTTGTAACCTCCACACACACACAGCCT 2061
Db 803 GlyIysMetArgGlySerThrGlnThrValGlyIleAlaAlaLysThrGlyGluAsnThr 822
QY 2062 CCTGCGCTAATCAGTTGATCAGCTTCCCAATCTGCATTTGTTCTTCTTCTTGTGTA 2121
Db 823 ThrAlaAla-----AlaThrLeuGlyMetGlyArgSerThrTrpSer 836
QY 2122 GCAACAATGAGTTCAGATCCTCTACCAATCTCCAGCGCAAGATTCATCTCTGCA 2181
Db 837 GluAsnSerAla-----AsnAlaLysThr-----AspSerIleSerLeu 849
QY 2182 GTCATTGCT--AGCACAACTGCTGTTCTGTTCACAAATAGTGGGCCCTATCTTTTTCAG 2238
Db 850 PheAlaGlyIleArgHisAspAlaGlyAspIleGlyTyrLeuLysGlyLeuPhe----- 867
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGAT-----TGCGTAGGTTCT 2283
Db 868 -----SerTyrGlyArgTyrLysAsnSerIleSerArgSerThrGlyAla 882
QY 2284 AATCAA-----AAATCAATGTCCTGAAATTTACAGTTAGGACTAAGCCCCCA 2331
Db 883 AspGluHisAlaGluGlySerValAsnGlyThrLeuMetGlnLeuGlyAlaLeuGlyGly 902
QY 2332 GCTAATGCCCA-----TCAGATTGACTCTAGGAAT----- 2364
Db 903 ValAsnValProPheAlaAlaThrGlyAspLeuThrValGluGlyGlyLeuArgTyrAsp 922
QY 2365 -----GAGATGCTCAAGTATCAAGAGCTGGAAGCTTTCGCGGATCCTAAT 2418
Db 923 LeuLeuLysGlnAspAlaPheAlaGluLysGlySer--AlaLeuGlyTrpSerGlyAsn 941
QY 2419 ACAGCAATAATGGT-----CCTATATCT 2442
Db 942 SerLeuThrGluGlyThrLeuValGlyLeuAlaGlyLeuLysLeuSerGlnProLeuSer 961
QY 2443 CTGAAGCTACATGACTAAACTGGGTATATCTCGGCCCTGAGCGAGTAGTCTCTTTG 2502
Db 962 AspyAlaValLeuPheAlaThrAla-----GlyValGluArgAspLeuAsnGly 978
QY 2503 GTTCCAAATAGTTATGGGGATCCATTTTAGATATACGATCTGCG----- 2547
Db 979 ArgAspTyrThrValThrGlyGlyPheThrGlyAlaThrAlaAlaThrGlyLysThrGly 998
QY 2548 -----CATTACGAATTCAGCAAGTGGATGGCGGCTCTTATTGTGCA 2592
Db 999 AlaArgAsnMetProHisThrArgLeuValAlaGlyLeuGlyAlaAspValGluPheGly 1018
QY 2593 GGATTTATGGTTCTCGAGTTTCGAATTTCTCTAT-----CAT 2631
Db 1019 AsnGlyTrp---AsnGlyLeuAlaArgTyrSerTyrAlaGlySerLysGlnTyrGlyAsn 1037
QY 2632 GACCGCGATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTTATCTCTAGGA--- 2688
Db 1038 HisSerGlyArgValGlyValGlyTyrArgPheLeuAspGlyGlyGlyThrGlySer 1057
QY 2689 -----GCAACTCTCTACTTTGGATCATCGATGTTGGTGTAGCATTTACCAAGTA 2739
Db 1058 SerAspLeuAlaAsnAspSerPheIleArgGlnValLeuAspArgGlnHisPheGluPro 1077
QY 2740 TTTGCTAGA-----NCTAAAGATTATGATGTGTCGTTCCCAATCATCAT 2784
Db 1078 AspGlyLysTyrHisLeuPheGlySerArgGlyGluLeuAlaGluArgSerGlyHis--- 1096
QY 2785 GCTTGATGATCCGCTTTATCTATCTACCAAGCTTTTATGTGGATCCTATTGTTTC 2844
Db 1097 -----IleGlyLeuGlyLysIleGlnSerHisGln-----LeuGlyAsnLeuMetIle 1112
```

```
QY 2845 GGAGATGCTTTATCTCCGTGCTAGCTACGGGTTT-----GGGAAT 2883
Db 1113 GlnGlnAlaAlaIleLysGlyAsnIleGlyTyrIleValArgPheSerAspHisGlyHis 1132
QY 2884 CAGCATATGAAAACCTCATATACATTTGCGAGGAGAGCGAT----- 2925
Db 1133 GluValHisSerProPheAspAsnHisAlaSerHisSerAspSerAspGluAlaGlySer 1152
QY 2926 -----GTTCGTTGGGATAAT----- 2940
Db 1153 ProValAspGlyPheSerLeuTyrArgIleHisTrpAspGlyTyrGluHisPheProAla 1172
QY 2941 AACTGCTCGGCTGAGAGATTGGAGCGGATTACCGATTGCGATTGATTCTCCATCTAAGCTC 3000
Db 1173 AspGlyTyrAspGlyProGlnGlyGlyTyrProAlaProLysGlyAlaArgAspIle 1192
QY 3001 TATTGTAATGATGCTGCTCTTCTCGTCAAA---GCTGAGTTTCTTATGCCGATCATGAA 3057
Db 1193 TyrSerTyrAspIleLysGlyValAlaGlnAsnIleArgLeuAsnLeuThrAspAsnArg 1212
QY 3058 TCTTTTACAGAGGAGGCGATCAAGCTCGGCGATTCAAGCGGACATCTCTCTAAATCTA 3117
Db 1213 Ser---ThrGlyGlnArgLeuAlaAspArgPheHisAsnAlaGlySerMetLeuThrGln 1231
QY 3118 TCAGTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTAAATAATATAGC 3177
Db 1232 GlyValGlyAspGlyPheLys-----ArgAlaThrArgTyrSer 1244
QY 3178 TTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGTGTTACTGAGACACGCTC 3237
Db 1245 ProGluLeuAspArgSerGlyAsnAlaAlaGluAlaPheAsnGlyThr----- 1260
QY 3238 CTATCCCATCAAGACATGACGACAAACAGATCCCTTTTCATTTAGCAAGACATGGAGTTGTG 3297
Db 1260 ----- 1260
QY 3298 GTTAGAGGATCTATGTTATGCTTCTCTAACAAGTAATATA-----GAAGTA 3342
Db 1261 -----AlaAspIleValLysAsnIleIleGlyAlaAlaGlyGluIle 1274
QY 3343 TATGCCCATGGAAGATATGAGTATCGAGATCTCTCGAGGCTATGGTTTTCAGTGCAGGA 3402
Db 1275 ValGlyAlaGly-----AspAlaValGln-----GlyIleSerGluGly 1287
QY 3403 AGTAGAGTC 3411
Db 1288 SerAsnIle 1290
```

RESULT 38

```
US-10-467-657-1088
; Sequence 1088, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: PONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1088
; LENGTH: 1468
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1088
```

Alignment Scores:

Db 304 SerHisIleSerSerAlaGlnGlyGlyLeuValThrSerSerSerAspAlaThrIle 323
QY 1192 AGTTTTCCAGAAATACGCTGCGTAGAGTTTGATGGGAACGTAGCCGAGTAGGAGGG 1251
Db 324 AsnPheSerGlyThrAlaAlaGlnArg-----AsnSer 334
QY 1252 ATTTACTCTACGGGAACGTTGCTTTCTGTAATAATGGAACCTTTGTTCTCAACAAT 1311
Db 335 IlePheSerGlyGlySerThrGlyAlaSerAlaGlnThrAlaThrAlaValIleAsn--- 353
QY 1312 GTTGCTTCTCTGTTTACATTGCTGCTAAGCAACCAACCAAGTGGACAGGCTTTCTAATACG 1371
Db 354 -----MetGlnAsnThr 357
QY 1372 AGTAATAATTACGGAGATCGAGAGCTATCTTCTGTAAGANTGGTGGCAAGCAGGATCC 1431
Db 358 AspIleThrValAspArgAsnGlySerLeuAlaLeuGlyLeuTrpAlaLeuSerGlyGly 377
QY 1432 AATAACTCTGGA---TCAGTTTCTTTGATGGAGGGAGTAGTTTCTTTAGTAGCAAT 1488
Db 378 ArgIleThrGlyAspSerLeuAlaIleThrGly----- 388
QY 1489 GTAGCTGTGGAAAGGGGAGCTATTATGCC----- 1521
Db 389 ---AlaAlaGlyAlaArgGly---IleTyrAlaMetThrAsnSerGlnIleAspLeuThr 406
QY 1522 AAAAGCTCTCGGTGCTAACTGTCGCCCTGTACAAATTTTAAAGGAATATCGCTAATGAT 1581
Db 407 SerAspLeuValIleAspMetSerThrProAspGlnMetAlaIleAlaThrGlnHisAsp 426
QY 1582 GTGGA-----GCGATTATTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATAT 1635
Db 427 AspGlyTyrAlaAlaSerArgIleAsnAlaSerGlyArgMetLeuIleAsn----- 443
QY 1636 CGAGATATTATTTCAGTGGGAATCTTAAAGA----- 1668
Db 444 GlySerValLeuSerIleGlyGlyLeuIleAsnLeuAspMethIleProGlySerValTrp 463
QY 1669 ACAGCCAAAGAGAACTGCTGCCGATGTTAATGCG-----GTAAGTGTGCTCTCA 1716
Db 464 ThrGlySerSerLeuSerAspAsnValAsnGlyGlyLysLeuAspValAlaMetAsnAsn 483
QY 1717 CAAGCCATTTCGATCGGATCGGAGGAGAAATAACGACATTAAGA----- 1761
Db 484 SerValTrpAsnValThrSerAsnSerAsnLeuAspThrLeuAlaLeuSerHisSerThr 503
QY 1762 -----GCTAAAGCAGGCGCATCAGATTCTCTTAATGATCCCATC 1800
Db 504 ValAspPheAlaSerHisGlySerThrAlaGlyThrPheThrThrLeuAsn-----Val 521
QY 1801 GAGATGGCAACCGAAATAACACGCGCAGCTCTTCCAAACTTCTTAAATAATACGAT 1860
Db 522 GluAsnLeuSerGlyAsnSerThrPheIleMetArgAlaAspValValGlyGluGlyAsn 541
QY 1861 GGTGAAGGATACACAGGAGATATCTTTGCTAATGAGACGACTACTTTGTACCAAAAT 1920
Db 542 GlyValAsnAsnArgGlyAspLeuLeuAsnIleSerGlySerAlaGlyAsnHisVal 561
QY 1921 GTTACGATACAGCAA-----GGAAGGATGTTCTTTCGT----- 1953
Db 562 LeuAlaIleArgAsnGlnGlySerGluAlaThrThrGlyAsnGluValLeuThrValVal 581
QY 1954 -----GAAAAGGCAAAATATCAGTGAATTTCTTAAGTACGACAGGTGGAGT 2001
Db 582 LysThrThrAspGlyAlaAlaSerPheSerAlaSerGlnValGluLeuGlyGlyTyr 601
QY 2002 CTGTAT-----ATGGAGCTGGAGTACATGGATTTGTA----- 2037
Db 602 LeuTyrAspValArgLysAsnGlyThrAsnTrpGluLeuTyrAlaSerGlyThrValPro 621
QY 2038 -----ACTCCACACCA-----CCACAACGCTCTCGCGCTAATCAG 2076
Db 622 GluProThrProAsnProGluProThrProAlaProAlaGlnProPro----- 637

QY 2077 TTGATCAGCTTTTCCAAATCTGCATTTGTCTCTTCTTCTTTGTAGCAACAATGCAGTT 2136
Db 638 -----1le 638
QY 2137 ACGAATCT---CCTPACCAATCTCCAGCGCAAGATTCATCTCTGCGATCATTTGTAGC 2193
Db 639 ValAsnProAspProThrProGluProAlaProThrProLysPro-----Thr 654
QY 2194 ACAACTGCT---GGTCTCTTACAATTAGTGGGCTTATCTTT----- 2232
Db 655 ThrThrAlaAspAlaGlyGlyAsnTyrLeuAsnValGlyTyrLeuLeuAsnTyrValGlu 674
QY 2233 -----TTTGAGGATTTGGATGATACAGCTTTATGATAGTATGAT 2271
Db 675 AsnArgThrLeuMetGlnArgMetGlyAspLeuArgAsnGlnSerLysAspGlyAsnIle 694
QY 2272 TGGCTA----- 2277
Db 695 TrpLeuArgSerTyrGlyGlySerLeuAspSerPheAlaSerGlyLysLeuSerGlyPhe 714
QY 2278 -----GGTCTTAATCAAAAAATCAATGTCTCTGAAA 2307
Db 715 AspMetGlyTyrSerGlyIleGlnPheGlyAspLysArgLysLeuSerAspValMetPro 734
QY 2308 TTACAGTTAGGACTAAG---CCCCCAGCTAATGCCCATCAGATTGTGACTCTAGGAAAT 2364
Db 735 LeuTyrValGlyLeuTyrIleAspSerThrHisAlaSerProAspTyrSerGlyGlyAsp 754
QY 2365 GAGATGCCCTPAAGTATGGCTATCAAGAGACCTGGAAGCTGCGTGGGATCCTAATACAGCA 2424
Db 755 GlyThrAlaArgSerAspTyrMetGlyMetTyrAla-----SerTyrMetAla 770
QY 2425 AATAATGGCTCTTACT-----CTGAAGCTACATGGACTAAAGTGGTATAAT 2475
Db 771 GlnAsnGlyPheTyrSerAspLeuValIleLysAlaSerArgGlnLysAsnSerPheHis 790
QY 2476 CTGGGCTGAGCGAGTAGTCTTTGTTGTTCCAAATAGTTTATGGGATCCATTTTAGAT 2535
Db 791 -----ValLeuAsp 793
QY 2536 ATACGATCTGCGCATTCAGCAATTCAGCAAGTGTGGTGGGCTCTTATTGT----- 2589
Db 794 SerGlnAsnAsnGlyValAlaAsnAlaAsnGlyThrAlaAsnGlyMetSerIleSerLeuGlu 813
QY 2590 -----CGAGGATATGGTCTTCTGGAGTTTCG 2616
Db 814 AlaGlyGlnArgPheAsnLeuSerProThrGlyTyrGlyPheTyrIleGluProGlnThr 833
QY 2617 AATTTCTCTATCATGACCGCATGCTTTTAGGTCAGGATATCGGTATATTATGCGGGT 2676
Db 834 GlnLeuThrTyrSerHisGlnAsnGluMetAlaMet-----LysAlaSerAsnGly 850
QY 2677 TATTCCTTAGGACCAAACTCTACTTTGGATCATCGATGTTGGTCTAGCATTTACCGAA 2736
Db 851 LeuAsnIleHisLeuAsnHisTyr-----GluSerLeuLeuGlyArgAla---SerMet 867
QY 2737 GTATTTGTAGATCTAAAGATTTAGTGTGCTGCTCCATCATCATGCTTCCATAGGA 2796
Db 868 IleLeuGly-----TyrAspIleThrAlaGlyAsnSerGlnLeu----- 880
QY 2797 TCCGTTTATCTATCTACCCAA-----CAAGCTTTATGTGA-----TCTATTGTTTC 2844
Db 881 AsnValTyrValLysThrGlyAlaIleArgGluPheSerGlyAspThrGluTyrLeuLeu 900
QY 2845 GAGATGCGCTTATCCGTGCTAGCTACGGGTTT---GGGAAT----- 2883
Db 901 AsnAspSer-----ArgGluLysTyrSerPheLysGlyAsnGlyTyrAsnAsnGlyVal 918
QY 2884 -----CAGCATATGAAACCTCATATACATTTCCAGAGGAGACGATGTCGT 2931
Db 919 GlyValSerAlaGlnTyrAsnLys---GlnHisThrPheTyrLeuGluAlaAspTyrThr 937

```

Qy 2932 TGGGATAAT 2940
Db 938 GlnGlyAen 940
RESULT 40
US-11-067-260-24
; Sequence 24, Application US/11067260
; Publication No. US20060051840A1
; GENERAL INFORMATION:
; APPLICANT: Arico, Maria
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISSERIAL PROTEINS
; FILE REFERENCE: 223002099501
; CURRENT APPLICATION NUMBER: US/11/067,260
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: 10/220,480
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: PCT/IB01/00420
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1392
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: deletion G983-961c
US-11-067-260-24
Alignment Scores:
Pred. No.: 9e-08 Length: 1392
Score: 190.00 Matches: 211
Percent Similarity: 32.9% Conservative: 120
Best Local Similarity: 21.0% Mismatches: 378
Query Match: 2.4% Indels: 296
DB: 11 Gaps: 52
US-10-701-844-1 (1-4435) x US-11-067-260-24 (1-1392)
Qy 559 TTTTCTCGAGAGAGTAACTTAAATAAATCTTGACAAATCTTATTCGACGCTTTGCTTTA 618
Db 268 TyrSerGlyAspLysThr-----AspGluGlyLeuArgLeuMetGlnGln 283
Qy 619 AGTTGTTTGGAACTTATAGGAGTTTACTGTTTATAGGAGGAGGACACTCGTTGACT 678
Db 284 SerAspTyrGlyAsnLeu-----SerTyrHisLeuArgAsnLysAsnMetLeuPheIle 301
Qy 679 TTC-----GAGNACATACGACTTCTACAAATGGGCGGAGCTCTA----- 717
Db 302 PheSerThrGlyAsnAspAlaGlnAlaGlnProAsnThrTyrAlaLeuLeuProPheTyr 321
Qy 718 AGTAATAGCGCTGCTGATGACTGTTTACTATTGAGGGT----- 756
Db 322 GluLysAspAlaGlnLysGlyLeuIleThrValAlaGlyValAspArgSerGlyGluLys 341
Qy 757 TTTAAGAAGATTATCCCTT-----TCCAATTGC 783
Db 342 PheLysArgGluMetTyrGlyGluProGlyThrGluProLeuGluTyrGlySerAsnHis 361
Qy 784 AATTCAATTACTTCGCTGCTGCTGCAACGACTTAATAAGGTAGCCAGACTCCGACG 843
Db 362 CysGlyIleThrAlaMetTrpCysLeuSerAlaProTyrGluAlaSerValArgPheThr 381
Qy 844 ACAACATCT-----ACACCGCTTAATGGTACTATTATTTCTAAACA 885
Db 382 ArgThrAsnProIleGlnIleAlaGlyThrSerPheSerAlaProIleValThrGlyThr 401
Qy 886 GATCTTTTCTTACTCAATAATGAGAAGTTCTGATTTCTATAGTAATTTACTCTCGAGAT 945
Db 402 AlaAlaLeuLeuLeu-----GlnLysTyrProTrpMetSerAsn----- 414
Qy 946 GGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTTC 1005
Db 415 -----AspAsnLeuArgThrThr 420

```

```

Qy 1006 CAAGAAAATACTGCTCAAGCTGATGGGGAGCTGTGTCAAGTAGTCACCAAGTTTC----- 1059
Db 421 LeuLeuThrThrAlaGln---AspIleGlyAlaValGlyValAspSerLysPheGlyTrp 439
Qy 1060 -----TCTGCTATGCTAAACGAGGCTCTCTATTGCTGTTTGTAGCGAAT 1101
Db 440 GlyLeuLeuAspAlaGlyLysAlaMetAsnGlyProAlaSerPheProPheGlyAspPhe 459
Qy 1102 GTTCGAGGAGTAAGAGGGGAGGAGGATGCTGCTGCTTCTAGGATGGCGACGAGGAGTGTC 1161
Db 460 ThrAlaAspThrLysGly----- 465
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTCAGAAAATACATCGCGTAGAGTTT 1221
Db 466 -----ThrSerAspIleAlaTyrSerPhe---ArgAsnAspIleSerGlyThr 480
Qy 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGGAACGTTGCTTTC--- 1278
Db 481 GlyGlyLeuIleLysGlyGlySerGlnLeuGlnLeuHisGlyAsnAsnThrTyrThr 500
Qy 1279 -----CTCAATAATATGGAACCTTGTCTTCTCAACAATGTTGCTTCTCTGTT 1326
Db 501 GlyLysThrIleIleGluGlyGlySerLeuValLeuTyrGlyAsnAsnLysSerAspMet 520
Qy 1327 TACATTGCTGCTAAGCAACCAACA-----AGTGGACAGGCTTCTTAATACAGTAATAAT 1380
Db 521 ArgValGluThrLysGlyAlaLeuIleTyrAsnGlyAlaAlaSerGlyGlySerLeuAsn 540
Qy 1441 -----GGATCATGTTTCTTTCATCGAGGAGGAGTAGTGTCTTCTTAGTACCAAT 1488
Db 559 ValHisIleLysGlySerLeuGlnLeuAspLysGlyThrLeuTyr----- 574
Qy 1489 GTAGCTGCTGGAAA-----GGGGAGCTATTATAT 1518
Db 575 ThrArgLeuGlyLysLeuLeuLysValAspGlyThrAlaIleIleGlyGlyLysLeuTyr 594
Qy 1519 GCCAAAAAG----- 1527
Db 595 MetSerAlaArgGlyLysGlyAlaGlyTyrLeuAsnSerThrGlyArgValProPhe 614
Qy 1528 CTCTCGGTGCTAACTGTGGC---CCTGTACAATTTTAAAGGAATATCGCTAATGATGAT 1584
Db 615 LeuSerAlaAlaLysIleGlyGlnAspTyrSerPhePheThrAsnIleGluThrAspGly 634
Qy 1585 GGAGCGATTTATTAGGAGAACTCGAGAGCTCAGTTTATCTGCTGATTATGGAGAT--- 1641
Db 635 GlyLeuLeuAlaSerLeuAspSerValGluLysThrAlaGlySerGlu---GlyAspThr 653
Qy 1642 -----ATTATTTTCGATGGGAATCTTAAAGAACAGCCAAA----- 1677
Db 654 LeuSerTyrTyrValArgArgGlyAsnAlaAlaArgThrAlaSerAlaAlaAlaHisSer 673
Qy 1678 -----GAGAATGCTCCGATGTTAATGGCTAACTGTGTCTCTCAACAGCC 1722
Db 674 AlaProAlaGlyLeuLysHisAlaValGluGlnGlyGlySerAsnLeuLeuLeuMet 693
Qy 1723 ATTTTCGATGGGA-----TCGGAGGGGAAAATAACAGCATTAAGAGCTTAAGCAGGG 1773
Db 694 ValGluLeuAspAlaSerGluSerAlaThrProGluThrValGluThrAlaAlaAla 713
Qy 1774 CATCAGATCTCTTAATGATCCCATCGAGATGGCAACGGGAATAACACGACGCGCAG 1833
Db 714 -----AspArgThrAspMetProGlyIleArgProTyrGlyAlaThr 727
Qy 1834 TCTTCCAAACTTCTAAAATAATTAACGATGTTGAAGGATACACAGGGGATATGTTTGTCT 1893
Db 728 PheArgAlaAlaAlaValGlnHisAlaAsnAlaAlaAspGlyValArgIlePheAsn 747

```

```
QY 1894 AATGAGCAGTACTTCTTCCAAATGTTAGATAGAG-----CAAGGAAGG 1941
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 748 SerLeuAlaAlaThrValTyrAlaAspSerThrAlaAlaHisAlaAspMetGlnGlyArg 767
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1942 ATTGTTCTTCGTGAAGCAAAATATCATAGTAATCTCTAAGTCACACAGGTGGAGT 2001
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 768 -----ArgLeuLysAlaValSerAspGlyLeuAspHisAsnGlyThrGly 782
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2002 CTGTATATGGAGCT-----GGAGTACATGGAT----- 2031
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 783 LeuArgValIleAlaGlnThrGlnAspGlyGlyThrTrpGlnGlnGlyValGlu 802
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2032 -----TTTGTAACCTCCACAACCAACCAACCAACCAAGCCT 2061
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 803 GlyLysMetArgGlySerThrGlnThrValGlyIleAlaAlaLysThrGlyGluAsnThr 822
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2062 CCTGCCGTAAATCAGTTGATCAGCTTCCAAATCTGCATTTGCTCTCTTCTCTTGTGA 2121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 823 ThrAlaAla-----AlaThrLeuGlyMetGlyArgSerThrTrpSer 836
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2122 GCACAATGAGTTACGATCTCTTACCATCTCCAGCCCAAGATCTCATCTGCA 2181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 837 GluAsnSerAla-----AsnAlaLysThr-----AspSerIleSerLeu 849
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2182 GTCATTGGT---AGCACAACTCTGTTCTGTGTACAAATTAGTGGCGCTATCTTTTGTAG 2238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 850 PheAlaGlyIleArgHisAspAlaGlyAspIleGlyTyrLeuLysGlyLeuPhe----- 867
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2239 GATTTGGATGATACATGATGATAGGTATGAT-----TGGCTAGGTTCT 2283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 868 -----SerTyrGlyArgTyrLysAsnSerIleSerArgSerThrGlyAla 882
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2284 AATCAA-----AAATCAATGTCCTGAAATTCACAGTTAGGAGTAAAGCCCCA 2331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 883 AspGluHisAlaGluGlySerValAsnGlyThrLeuMetGlnLeuGlyAlaLeuGlyGly 902
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2332 GCTAATGCCCA-----TCAGATTGACTCTAGGGAAT----- 2364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 903 ValAsnValProPheAlaAlaThrGlyAspLeuThrValGluGlyLeuArgTyrAsp 922
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2365 -----GAGATCGCTAAGTATGCTATCAAGGAAGCTTGGAGCTTGGCGGATCCTAAT 2418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 923 LeuLeuLysGlnAspAlaPheAlaGluLysGlySer---AlaLeuGlyTrpSerGlyAsn 941
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2419 ACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGAGCTAAAGCTGGTATATCTCT 2478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 942 SerLeuThrGluGly----- 946
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2479 GGGCCTGACGAGTAGCTTCTTTGTTCCAAATAGTTTATGGGATCCATTTTAGAT--- 2535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 947 -----ThrLeuValGlyLeuAlaGlyLeuLysLeuSerGlnProLeuSerAspLys 963
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2536 ---ATACGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGA 2592
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 964 AlaValLeuPheAlaThrAlaGlyValGluArgAspLeuAsnGlyArgAspTyrThr--- 982
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2593 GGATATGGGTTTCGAGTTTCGAATTTCTCTATCATGACCGCGATGCTTTAGGTCAG 2652
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 983 -----ValThrGlyGlyPheThrGlyAlaThrAlaAlaThrGlyLys 996
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2653 -----GGATATCGGTATTTAGTGGGGTTATTCTTCTTAGGAGCA 2691
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 997 ThrGlyAlaArgAsnMetProHisThrArgLeuValAlaGly-----LeuGlyAla 1013
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2692 AACTCTACTTTTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1014 AspValGluPheGlyAsnGlyTyrAsnGlyLeuAlaArgTyrSerTyrAlaGly---Ser 1032
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2752 AAAGATTATGTAGTGTGCTTCCAAATCATCATGCTTGGCATAGGATCCGTTTATCTATCT 2811
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1033 LysGlnTyr-----GlyAsnHisSerGlyArgValGlyValGlyTyr----- 1046
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2812 ACCCAACAAGCTTTATGTGGATCCTATTTGTTCCGGAGATCG----- 2853
```

```
Db 1047 -----ArgPheLeuGluGlyGlyGlyThrGlySerAlaThrAsnAspAspVal 1064
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2854 -----TTTATCCGTCTAGCTACGGTTTGGGAATCAG----- 2886
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1065 LysLysAlaAlaThrValAlaIleAlaAlaIleAlaIleAlaIleAlaIleAsnGly 1084
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2887 -----CATATGAAACCTCATATACATTTTCAGAGGAGAGCGATGTTCTGTTGGGATAAT 2940
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1085 PheLysAlaGlyGluThrIleTyrAspIleAspGluAspGlyThrIleThr---LysLys 1103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2941 AACTGCTCGCTCGAGAGATT-----GGAGCGGATTTACCGATTGTC 2982
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1104 AspAlaThrAlaAlaAspValGluAlaAspPheLysGlyLeuGlyLeuLysVal 1123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2983 ATTACTCCATCTAAGCTCTATTGTAATGAGTTGCGTCTCTTTCGTCGAAGCTGAGTTTCT 3042
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1124 ValThrAsnLeuThrLysThrValAsnGluAsnLysGlnAsnValAspAlaLysValLys 1143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 3043 TATGCCGATCATGAA 3057
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1144 AlaAlaGluSerGlu 1148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: May 13, 2006, 12:17:46
Job time : 621.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 13, 2006, 10:00:08 ; Search time 28 Seconds

(without alignments)
4572.014 Million cell updates/sec

Title: US-10-701-844-1

Perfect score: 7883

Sequence: 1 999caaaactctcccccg.....gcactctaacgaagaattc 4435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DRV=xlp
-O=/abes/ABSSWEB spool/US10701844/runat 12052006 165423 26326/app query.fasta_1
-DB=PIR -QFMT=faetan -SUFFIX=trp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bic -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p
-USER=US10701844 @CNG 1.1.63 @runat 12052006 165423 26326 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DSELP=6 -DELEXT=7

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5131.5	65.1	1013	2	G71460
2	3846	48.8	987	2	H81722
3	1589	20.2	1016	2	H71460
4	1366	17.3	995	2	C81593
5	1365.5	17.3	973	2	B86547
6	1365.5	17.3	973	2	F72076
7	1179	15.0	983	2	A81723
8	1101	14.0	928	2	G86546
9	1101	14.0	928	2	G81591
10	1092	13.9	949	2	F81591
11	1090	13.8	928	2	D72077
12	1090	13.8	928	2	H86546
13	1083	13.7	930	2	D86546
14	1083	13.7	930	2	A81591

15	1079	13.7	930	2	D72078	polymorphic outer
16	1041	13.2	936	2	C72078	polymorphic outer
17	1040	13.2	936	2	B81591	polymorphic membra
18	1040	13.2	936	2	C86546	polymorphic outer
19	1021	13.0	928	2	B72077	polymorphic membra
20	1021	13.0	928	2	E86546	polymorphic outer
21	1015	12.9	1276	2	B86546	polymorphic outer
22	1015	12.9	1276	2	C81591	polymorphic membra
23	991.5	12.6	1407	2	B72078	polymorphic outer
24	979.5	12.4	772	2	H86492	Pmp 3 [imported] -
25	973	12.3	922	2	B72131	polymorphic outer
26	973	12.3	922	2	E86491	polymorphic outer
27	972	12.3	922	2	F81539	polymorphic membra
28	939.5	11.9	841	2	F72130	polymorphic membra
29	789	10.0	867	2	B81721	polymorphic membra
30	771.5	9.8	878	2	B71460	probable outer mem
31	703	8.9	712	2	E86492	polymorphic outer
32	702	8.9	978	2	G72076	polymorphic outer
33	697	8.8	978	2	C86547	polymorphic membra
34	697	8.8	978	2	B81593	polymorphic membra
35	679.5	8.6	976	2	F81722	polymorphic membra
36	649.5	8.2	964	2	E71460	probable outer mem
37	644	8.2	938	2	F86548	polymorphic outer
38	644	8.2	938	2	H72074	polymorphic membra
39	611	7.8	1520	2	A81731	polymorphic outer
40	609	7.7	946	2	C86549	polymorphic membra
41	609	7.7	946	2	D81594	polymorphic membra
42	607	7.7	946	2	C72075	polymorphic outer
43	599.5	7.6	947	2	G86557	polymorphic membra
44	599.5	7.6	947	2	D72067	polymorphic membra
45	597.5	7.6	1531	2	H71468	probable outer mem

ALIGNMENTS

RESULT 1

G71460
probable outer membrane protein G - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C/Species: Chlamydia trachomatis
C/Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C/Accession: G71460
R/Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell Science 282, 754-759, 1998
A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra A/Reference number: A71570; MUID:99000809; PMID:9784136
A/Accession: G71460
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1013 <ARN>
A/Cross-references: UNIPROT:O84879; UNIPARC:UPI0000131CF5; GB:AE001360; GB:AE001273; NI A/Experimental source: serotype D, strain UW-3/Cx
C/Genetics:
A/Gene: pmpG

Alignment Scores:
Pred. No.: 0
Score: 5131.50
Percent Similarity: 98.8%
Best Local Similarity: 97.0%
Query Match: 65.1%
DB: 2
Length: 1013
Matches: 983
Conservative: 18
Mismatch: 11
Indels: 1
Gaps: 1

US-10-701-844-1 (1-4435) x G71460 (1-1013)

QY	382	ATGCAACGCTCTTTCCATAGTTCTTTCTTCAATGATTCTAGCTATTCTTCTGCTCT	441
DB	1	MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaIleSerCysCys	20
QY	442	TTAAATGGGGGGGATATGACAGCAAAATCATGTTCTCTCAAGGAATTTACGATGGGAG	501
DB	21	LeuSerGlyGlyGlyTyraAlaAlaGluIleMetIleProGlnGlyIleTyraAspGlyGlu	40
QY	502	ACGTTAACTGATCATCTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTT	561

Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTGAGGAGAGTTAAACATTAATAAATCTTGACAAATCTATTGACAGCTTTGCTTTAAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
Qy 622 TGTGTTGGAACTTATTATGAGGAGTTTACTGTTTGGGAGAGGAGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACGAGCTTCTACAAATGGCGAGCTCTAACTAATAGCGCTGCTGATGAGCTG 741
Db 101 GluAsnIleLeuThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTACTATTGAGGGTTTTAAAGAAATATCTTTTCCAAATTTGCAATTTCAATTTACTTCCGCTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTGCTGCTGCAACGACTAATAAGGTAGCGAGCTCCGAGACGACAACTCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
Qy 862 AATGCTACTATTATTCTTAAACAGATCTTTTGTACTCAATAATGACGAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
Qy 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluLeuLeuThrAlaGlnAlaAspGlyLysAlaCys 220
Qy 1042 CAAAGTACCAAGCTTCTCTGCTAGCTTAAGCTTAAGAGGCTCTATTGCTTGTACCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy 1102 GTTGAGAGATGAAGGGGAGGAGTGTCTGCTGTTTCAGGATGGCGAGGAGGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyLysAlaValAlaGlnAspGlyGlnGlnGlyValSer 260
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Qy 1222 GATGGGAAGTACCGGAGTAGGAGGAGGATTTACTCTCTACGGGAACGTTGCTTTCCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Qy 1282 AATAATGGAACCTTGTCTCAACATGTTGCTTCTCTGTTTACATTTGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu 320
Qy 1342 CAACCAACAAGTGGACAGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
Qy 1402 TTTCTGTAGAATGGTGGCAA---CGAGATCCAATACTCTGGATCAGTTTCTTTGAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaLysSerAsnAsnSerGlySerValSerPheAsp 360
Qy 1459 GGAGAGGAGTATGTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380
Qy 1519 GCCAAAGAGCTCTCGGTGCTAATCTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
Qy 1579 GATGCTGGAGCGATTTATTAGAGAACTCTGGAGAGCTCAGTTTATCTGCTCATTTATGGA 1638

Db 401 AspGlyGlyAlaIleTyrLeuGlyLysSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAAATGCTGCCGATGTTAAT 1698
Db 421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
Qy 1699 GGCTTAATCTGCTCCTCACAGCCATTTCCAGTGGATCGGGAGGAAATAACACACATT 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
Qy 1759 AGACTAAAGCAGGCGCATCAGATTTCTTTAAATCATCCCATCGAGATGGCAACCGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
Qy 1819 AACGACGACGCGAGCTTCTCCAACTTCTTAAATTAACGATGTGTGAAGGATACACAGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500
Qy 1879 GATATTCTTTTTCGTAATGGAGCAGTACTTTGTACCAAAATGTTTACGATAGACCAAGGA 1938
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Qy 1939 AGGATTGTTCTTCTGAAAAGGCCAAATTAATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Qy 1999 AGTCTGATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACCAACCAACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
Qy 2059 CCTCTCGCGCTAATCAGTTGATCAGCTTTCCTCAATCTGCAATTTGTTCTTTCTCTTTG 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
Qy 2119 TTAGCAACAATGCAGTTTACGAATCTCTACCAATCTCCAGGCGCAAGATTTCTCATCCT 2178
Db 581 LeuAlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
Qy 2179 GCAGTCAATGTTAGCACAACTGCTGTTTCTTTTCAATTAAGTGGGCTTATCTTTTTCAG 2238
Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy 2239 GATTTCGATGATCAGCTTATGATAGTATGATGCTTGGCTAGGTTCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
Qy 2299 GTCTGAAAATACAGTTAGGAGCTAAGCCCCCAGCTAATGCCCATCAGATTTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Qy 2359 GGGAAATGAGTGCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCGGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680
Qy 2419 ACAGCAAAATATGCTTACTTACTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700
Qy 2479 GGGCTGAGCAGTAGCTTCTTTGTTTCCAAATAGTTTATGGGATCCATCTTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720
Qy 2539 CGATCTGCGCATTCAGCAATTCAGGAAGTGTGGAATGGCGCTCTTATTTGCGAGGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy 2599 TGGGTTTCTGGAGTTTCGAATTTCTTCTCATCATGACCGGATGCTTTTAGGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy 2659 CGGTATATTAGTGGGGGTTTATTTCTTAGGAGCAATCTCTACTTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780

QY 2719 GGTCTAGCATTTACCGAAGTATTGGTAGATCTAAGATTTATGCTAGTGCTGTTCCAAAT 2778
 DB GlyLeuAlaPheThrGluValPheGlyArgSerLeuValValCysArgSerAsn 800
 QY 2779 CATCATGCTTGGCATAGATCCGTTTATCTATCTATCTACCCAAAGCTTTATGTGATCCTAT 2838
 DB HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
 QY 2839 TTGTTCCGAGATCGGTTTATCCGTGCTAGCTACCGGTTTGGGATCCGATCAGCATATGAACCC 2898
 DB LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
 QY 2899 TCATATACATTTCCAGAGAGACGATGTTGTTGGGTAATAACTGTCTGGCTGGAGAG 2958
 DB SerTyrThrPheAlaGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
 QY 2959 ATTGGAGCGGGATTACCGATTGTGATTACTCCATTAAGCTCTATTGTAATGAGTTGCGT 3018
 DB IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
 QY 3019 CCTTTCCGTCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGGAT 3078
 DB ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
 QY 3079 CAAGCTCGGGCATTCAAGCGGACATCTCCTAAATCTATCACTTCTGTTGGAGTGAAG 3138
 DB GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
 QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAAATATAGCTTTTATGCGCGCTTATATCTGT 3198
 DB PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
 QY 3199 GATGCTTATCGCAACCATCTCTGTATCTGACAGACAAACGCTCTCTATCCCATCAAGAGACATGG 3258
 DB AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
 QY 3259 ACAACAGATGCTTTCATTTAGCAACATGGAAGTTGTGGTTAGAGCATCTATGATGCT 3318
 DB ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 QY 3319 TCTCTAACAGTAAATATAGAAATATATGCGCATGGAAGATATGATATCCAGATGCTTCT 3378
 DB SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 QY 3379 CGAGGCTATGTTGATGTCAGGAAGTATGAGTCCGGTTC 3417
 DB ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
 RESULT 2
 H81722
 polymorphic membrane protein G family TC0263 [imported] - Chlamydia muridarum (strain N1
 C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C;Accession: H81722
 R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey,
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A;Reference number: A81500; MUID:20150255; PMID:10684935
 A;Accession: H81722
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-987 <TET>
 A;Cross-references: UNIPROT:Q9PL45; UNIPARC:UPI0000057867; GB:AE002293; GB:AE002160; NII
 A;Experimental source: strain Nigg (MoPn)
 C;Genetics:
 A;Gene: TC0263
 Alignment Scores:
 Pred. No.: 8,25e-257 Length: 987
 Score: 3845.00 Matches: 739
 Percent Similarity: 82.9% Conservative: 106

Best Local Similarity: 72.5% Mismatches: 136
 Query Match: 48.8% Indels: 38
 DB: 2 Gaps: 13
 US-10-701-844-1 (1-4435) x H81722 (1-987)
 QY 379 GTGATGCAAAACGCTTTTCCATAGTCTTTTCTTTCAATGATTTCTAGCTTATTTCTGCTGC 438
 DB 1 MetMetGlnThrProPheHisLysPhePheLeuLeuAlaMetLeuSerTyr ----- 17
 QY 439 TCTTTAAATGGGGGATATCGACAGAAATCATGGTTCCTCAAGGAATTTACGATGGG 498
 DB 18 SerLeuLeuGlnGlyGlyHisAlaAlaAspIleSerMetProGlyIleTyrAspGly 37
 QY 499 GAGACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTT 558
 DB 38 ThrThrLeuThrAlaProPheProTyrThrValIleGlyAspProArgGlyThrLysVal 57
 QY 559 TTTTCTGCGAGAGATTAAACATTAATAAATCTTGACATTTCTATTCGACGTTTGCCTTTA 618
 DB 58 ThrSerSerGlySerLeuGluLeuLysAsnLeuAspAsnSerIleAlaThrLeuProLeu 77
 QY 619 AGTTGTTTGGGAACCTTATAGGAGTTTTACTGTTTATAGGAGAGACACTCGTTGACT 678
 DB 78 SerCysPheGlyAsnLeuLeuGlyAsnPheThrIleAlaGlyArgGlyHisSerLeuVal 97
 QY 679 TTCGAGAACATACGACATTTCTCAAAATGGCGAGCTCTAAAGTAATAGCGCTGCTGATGGA 738
 DB 98 PheGluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnHisAlaProSerGly 117
 QY 739 CTGTTTACTATTGAGGGTTTTAAAGAAATTTATCCTTTTCCAATTTGCAATTCATTTACTTGC 798
 DB 118 LeuPheValIleGluAlaPheAspGluLeuSerLeuLeuAsnCysAsnSerLeuValSer 137
 QY 799 GTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACT---CCGACGACAACTCTACA 855
 DB 138 ValValPro-----GlnThrGlyGlyThrThrThrSerVal 149
 QY 856 CCGTCTAAGTACTATTATTCTTAAACAGACTTTTGTGTTACTCAATATAGAAAGTTC 915
 DB 150 ProSerAsnGlyThrIleTyrSerArgThrAspLeuValLeuArgPheIleLysVal 169
 QY 916 TCATTCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGACTTAACG 975
 DB 170 SerPheTyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaGlnSerLeuMet 189
 QY 976 GTTCAAGGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGA 1035
 DB 190 ValAsnGlyIleGluLysLeuCysThrPheGlnGluAsnValAlaGlnSerAspGlyGly 209
 QY 1036 GCTTGTCAAGTAGTACCAGTTTCTGCTATGCTTAAACGAGGCTCTTATGCTTTGTA 1095
 DB 210 AlaCysGlnValThrLysThrPheSerAlaValGlyAsnLysValProLeuSerPheLeu 229
 QY 1096 GCCAATGTTTCAGGAGTAAAGAGGGGAGGATGCTGCTTTCAGGATGGGCGAGCGGA 1155
 DB 230 GlyAsnValAlaGlyAsnLysGlyGlyValAlaAlaValLysAspGly---GlnGly 248
 QY 1156 GTGTCATCATCTACTTCAACAGAGAATCCAGTAGTAGTAAAGTTTTCAGAAATATCTCGGTA 1215
 DB 249 AlaGlyGlyAlaThr-----AspLeuSerValAsnPheAlaAsnAsnThrAlaVal 265
 QY 1216 GAGTTTCATGGGAACGTAGCCCGAGTAGGAGGAGGATTTTACTCCTACGCGGAGCTTGCT 1275
 DB 266 GluPheGluGlyAsnSerAlaArgIleGlyGlyIleTyrSerAspGlyAsnIleSer 285
 QY 1276 TTCCTGAATTAATGAAAAACCTTGTGTTCTCAACAATGTTCTCTCTCTCTGTTTACATT--- 1332
 DB 286 PheLeuGlyAsnAlaLysThrValPheLeuSerAsnValAlaSerProIleTyrValAsp 305
 QY 1333 ---GCTGCTAAGCAACCAACAGTGGACAGGCTTCTTAATACGATTAATATTACGGAGAT 1389
 DB 306 ProAlaAla-----AlaGlyGlyGlnProProAlaAspLysAspAsnTyrGlyAsp 322

```
QY 1390 GGAGGAGCTATCTCTGTGAAGATGGTGGCAAGCAGGATCCAATAACTCTGGATCAGTT 1449
Db 323 GlyAlaIlePheCysLeuAen-----AepThrAenIleGlyGluVal 337
QY 1450 TCCTTTGATGAGAGGAGTATTTCTTTAGTAGCAATGTAGCTGCGGAAGGGGGA 1509
Db 338 SerPheLysAepGluGlyValValPhePheSerLysAenIleAlaAlaGlyLysGly 357
QY 1510 GCTATTTATGCCAAAAGCTCTCGGTTCCTAACTGTGCCCTGTACAAATTTTAAAGGAAT 1569
Db 358 AlaIleTyrAlaLysLeuThrIleSerAepCysGlyProValGlnPheLeuGlyAen 377
QY 1570 ATCGCTAATGATGGGAGCGCATTTATTATGAGAAATCTGAGAGCTCAGTTTATCTGCT 1629
Db 378 ValAlaAenAepGlyAlaIleTyrLeuValAepGlnGlyLeuSerLeuSerAla 397
QY 1630 GATTATGAGATATTATTTTCGATGGGAATCTTAAAGACAGCCAAAGAGAAATGCTGCC 1689
Db 398 AepArgGlyAepIlePheAepGlyAenLeuLysArgMetAlaThrGlnGlyAlaAla 417
QY 1690 GATGTTAATGGCTAACTGTGCTCCACAAAGCCATTTGATGGGATCGGAGGAAATA 1749
Db 418 ThrValHisAepValMetValAlaSerAenAlaIleSerMetAlaThrGlyGlnIle 437
QY 1750 ACGACATTAAAGACTAAAGCGGCATCAGATTCTTTTAATGATCCCATCGAGATGGA 1809
Db 438 ThrThrLeuArgAlaLysGluGlyArgArgIleLeuPheAenAepProIleGluMetAla 457
QY 1810 AACGAAATACACGCGCGAGCTCTCCAACTTCTAAATAATTAAAGATGTCGAAGA 1869
Db 458 AenGly-----GlnProVal-----IleGlnThrLeuThrValAenGluGlyGlu 473
QY 1870 TACACAGGGATATTTGTTTCTTAATGAAGCAGTACTTTCTACAAATGTTTACGATA 1929
Db 474 TyrThrGlyAepIleValPheAlaLysGlyAepAenValLeuTyrSerSerIleGluLeu 493
QY 1930 GAGCAAGGAGGATGTTCTTCGTGAAGAGCAAAATTATCATGTAATCTCTAAAGTCAG 1989
Db 494 SerGlnGlyArgIleIleLeuArgGluGlnThrLysLeuValAenSerLeuThrGln 513
QY 1990 ACAGTGGAGTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACCA 2049
Db 514 ThrGlyGlySerValHisMetGluGlyGlySerThrLeuAepPheAlaVal----- 530
QY 2050 CCACAACAGCTCTCGCGCTAATCAGTTGATCAGCTTTCCAATCTGCATTTGCTCTT 2109
Db 531 ----ThrThrProProAlaAlaAenSerMet---AlaLeuThrAenValHisPheSerLeu 548
QY 2110 TCTTCTTTGTAGCAAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2169
Db 549 AlaSerLeuLeuLysAenAenGlyValThrAenProProThrAenProProValGlnVal 568
QY 2170 TCTCATCTCGAGTCAATGTTAGCACAACTGCTGTTCTGTACAAATGATGGGCTATC 2229
Db 569 SerSerProAlaValIleGlyAenThrAlaAlaGlyThrValThrIleSerGlyProIle 588
QY 2230 TTTTGTGAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2289
Db 589 PhePheGluAepLeuAepGluThrAlaTyrAepAenAenGlnThrLeuGlyAlaAepGln 608
QY 2290 AAAATCAATGCTCTGAAATTTACAGTTAGGAGCTTAAGCCCCCAGCTTAATCCCCATCAGAT 2349
Db 609 ThrIleAepValLeuGlnLeuHisLeuGlyAlaAenProProAlaAenAlaProThrAep 628
QY 2350 TTGACTCTAGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2409
Db 629 LeuThrLeuGlyAenGluSerSerLysTyrGlyTyrGlnGlySerThrLeuGlnThr 648
QY 2410 GATCCTTAATACACCAAT-----AATGGTCTTATATCTCTGAAAGCTACATCGACT 2460
Db 649 GluProAepProAlaAenProProGlnAenAenSerTyrMetLeuLysAlaSerThrThr 668
```

```
QY 2461 AAAACTGGGTATTAATCCTGGGCGCTGAGCAGTAGCTTCTTTGGTTCCAAATAGTTTATGG 2520
Db 669 LysThrGlyTyrAenProGlyProGluArgValAlaSerLeuValSerAenSerLeuTrp 688
QY 2521 GGATCCATTTTATGATATACGATCTGCGCATTTCAACAATTCGAAGTGTGTGATGGGCGC 2580
Db 689 GlySerIleLeuAepValArgSerAlaHisSerAlaIleGlnAlaSerIleAepGlyArg 708
QY 2581 TCTTATTGTCGAGATATTGGGTTTCTGGAGTTTTCGAATTTCTTCTATCATGACCGCAT 2640
Db 709 AlaTyrCysArgGlyIleTrpIleSerGlyIleSerAenPhePheTyrHisAepGlnAep 728
QY 2641 GCTTTAGTCAAGCATCGGTATATAGTGGGGTTTATCTCTTAGGAGCAAACTCCCTAC 2700
Db 729 AlaLeuGlyGlnGlyTyrArgHisIleSerGlyGlyTyrSerIleGlyAlaAenSerTyr 748
QY 2701 TTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATTGTTAGTATCTAAAGATTAT 2760
Db 749 PheGlySerSerMetPheGlyLeuAlaPheThrGluThrPheGlyArgSerLysAepTyr 768
QY 2761 GTAGTGTGCTGCTTCCATCATCTGCTGCTGATAGATCCGTTTATCTATCTACCCAA 2820
Db 769 ValValCysArgSerAenAepHisThrCysValGlySerValTyrLeuSerThrArgGln 788
QY 2821 GCTTTATGTGATCCTATTTGTTTCGAGATCGGTTTATCTCTGCTAGCTACGGGTTTGGG 2880
Db 789 AlaLeuCysGlySerCysLeuPheGlyAepAlaPheValArgAlaSerTyrGlyPheGly 808
QY 2881 AATCAGCATATGAAAACTCATATACATTTTCGAGAGGAGCGATGTTGTTGGGATAAT 2940
Db 809 AenGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAenValArgTyrAepAen 828
QY 2941 AACTGTCTGCTGAGAGATTTGGAGCGGATTCAGTATCCGATGTGATTCCTCATCTAAGCTC 3000
Db 829 AenCysValValGlyGluValGlyAlaGlyLeuProIleMetLeuAlaAlaSerLysLeu 848
QY 3001 TATTTGATGATGTTGCTGCTTCGTCGCAAGCTGTTTCTTATGCGCATCATGAACTCT 3060
Db 849 TyrLeuAenGluLeuArgProPheValGlnAlaGluPheAlaTyrAlaGluHisGluSer 868
QY 3061 TTTTACAGAGGAGCGCATCAAGCTCGGCGCATTTCAAGAGCGGACATCTCTAAATCTATCA 3120
Db 869 PheThrGluArgGlyAepGlnAlaArgGluPheLysSerGlyHisLeuMetAenLeuSer 888
QY 3121 GTTCTGTGTGAGTGAAGTTGATGATGTTCTAGTACACATCTCTTAATAATATAGCTTT 3180
Db 889 IleProValGlyValLysPheAepArgCysSerSerLysHisProAenLysTyrSerPhe 908
QY 3181 ATGGCGGCTTATCTGTGATGCTTATCGCACCATCTCTGTTACTGAGACAGCTCCTA 3240
Db 909 MetGlyAlaTyrIleCysAepAlaTyrArgSerIleSerGlyThrGluThrThrLeuLeu 928
QY 3241 TCCCATCAAGAGACATGAGCAACAGATCGCTTTTCAATTTTAGCAAGACATGGAGTTGGTT 3300
Db 929 SerHisLysGluThrThrThrAspAlaPheHisLeuAlaArgHisGlyValMetVal 948
QY 3301 AGAGATCTATGATGCTTCTCTAACAAAGTAAATATAGAAATATATAGCCATGGAAGATAT 3360
Db 949 ArgGlySerMetTyrAlaSerLeuThrGlyAenIleGluValTyrGlyHisGlyLysTyr 968
QY 3361 GAGTATCGAGATGCTTCTCGAGGCTATGTTTGTGTCGAGGAAGTAGAGTCGGGTTCC 3417
Db 969 GluTyrArgAepAlaSerArgGlyTyrGlyLeuSerIleGlySerLysIleArgPhe 987
```

RESULT 3

H71460
probable outer membrane protein H - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
R:Accession: H71460
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A>Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis

A;Reference number: A71570; MUID:99000809; PMID:97841136
A;Accession: H71460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1016 <ARN>
A;Cross-references: UNIPROT:O84880; UNIPARC:UPI0000046AC7; GB:AE001360; GB:AE001273; NID
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: pmph
C;Superfamily: Polymorphic membrane protein H family

Alignment Scores:		
Pred. No.:	2.3e-101	Length: 1016
Score:	1589.00	Matches: 311
Percent Similarity:	92.7%	Conservative: 5
Best Local Similarity:	91.2%	Mismatches: 10
Query Match:	20.2%	Indels: 15
Gap:	2	Gaps: 2

US-10-701-844-1 (1-4435) x H71460 (1-1016)

QY	3451	ATGCCTTTTCTCTTTGAGATCATCATATTTGTTTTTTAGCTTGTTGTTGTTCTATTTCG	3510
Db	1	MetProPheSerLeuAlaGSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer	20
QY	3511	TATGGATTCCGGAGCTCTCCTCAAGTGTTAAACGGCTAATGTAACCACTCTCTTTTAAAGGGA	3570
Db	21	TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrProPheLysGly	40
QY	3571	GACGATGTTTACTTGATGGAGACTGGCGCTTTCTCAATGCTATGCAGGAGCTGAAGAA	3630
Db	41	AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn	60
QY	3631	GGTTCGATTATCTCAGCTAATGGGCACAAATTAAAGATTACCGACACAAACCATACATTA	3690
Db	61	GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu	80
QY	3691	TCATTATCAGATTCTCAAGGCCAGTTCTTCCAAATATTGCTTCATTCTCAGCAGGAGAG	3750

QY	3751	ACACTTACTCTGAGAGATTTTTCAGAGTCTGATGTTCTCGAAAAATGTTTCTTCGCGGAA	3810
DB	101	ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu	120
QY	3811	AAGGAATGATCTCCGGGAAAACCTGGAGTATTTCCGAGACAGCGAGTGATGATTTTCTCG	3870
DB	121	LysGlyMetIleSerGlySerThrValSerIleSerGlyAlaGlyGluValIlePheTrp	140
QY	3871	GATAACTCCGTGGGGTATTCTCCTTTATCTACTGTGCCAAACCTCATCATCAACTCCGCCT	3930
DB	141	AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro	160
QY	3931	GCT-----CCACAGTTAGTCATGCTCGG	3954
DB	161	AlaProAlaProAlaProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArg	180
QY	3955	AAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGGCGTCAAAAAGGGGTC	4014
DB	181	LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysLysGlyVal	200
QY	4015	ATGTTTCGATAATAATGCCGGGAATTTCCGAAACAGTTTTTTCGAGGTAGAAGATAATAAT	4074
DB	201	MetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsn	220
QY	4075	GCTGTGTGTGGAGGCAGTGGGTTCCGCTACCATCAAGTACGACTTTTACAGTTTAAAAA	4134
DB	221	AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs	238
QY	4135	CTGTAAGGGAAGTTTCTTTTCACAGATAACGTACGCCTCTTCGCGAGCGGAGTGTTT	4194
DB	238	nCyLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyGlyValVal	258

Qy	4195	TAAAGGCATTTGCTTTTCCAAAGACAATGAAGAGGCGCATATCTTCGAGGGAACACAGC	425
Db	258	rLygLyThrValLeuPheLysAspGluGlyGlyLeuPheArgGlyAsnThrAl	278
Qy	4255	ATACGATGATTTAAGGATCTTGCTGCTACTAATACGAGTACGAATACGGAGACAGAGG	4314
Db	278	aTryAspAspLeuGlyLeuAlaThrSerArgAspGlnAsnThrGluThrGlyGly	298
Qy	4315	CGGTGAGGAGTTATTGCTCTCCAGATGATCTGTAAAGTTTGAAGGCCAATAAAGTTC	4374
Db	298	yGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlyLys	318
Qy	4375	TATTGTTTTTGATTACAACTTTCMAAAGGCAGAGCGGGAAGCATCTTAACGAAGAATT	4434
Db	318	rIleValPheAspTryAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh	338
Qy	4435	C 4435	
Db	338	e 338	

RESULT 4

C81593
polymorphic membrane protein G family CP0299 [imported] - Chlamydophila pneumoniae (str
C/Spectes: Chlamydophila pneumoniae, Chlamydia pneumoniae
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
C/Accession: C81593
R/Read, T.D.; Bruham, R.C.; Shen, C.; Gall, S.R.; Heidelberg, J.P.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
X:Reference number: 281500. MIMD:20150255. PMID:10684935

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-995 <REA>
A;Cross-references: UNIPARC:UPI00001655F9; GB:AE002191; GB:AE002161; NID:g7189216; PIDDN:
A;Environmental source: strain AR39. HT. Cells

A;Gene: CP0299

Alignment Scores:			
Pred. No.:	5,238-86	Length:	995
Score:	1366.00	Matches:	356
Percent Similarity:	49.4%	Conservative:	172
Best Local Similarity:	33.3%	Mismatches:	425
Query Match:	17.3%	Indels:	116
DB:	2	Gaps:	29

US-10-701-844-1 (1-4435) x C81593 (1-995)

QY	331	TTAATTTTTTAAGACGGGAGCTAAATAATTTTATCTCTC-----	369
Db	3	LeuPhePheTySer-----LeuSerLeuLeuCysArgIleileTrpPhe	17
QY	370	--AGCTTTTGTCATGCCAAACGTCTTCCATAGTTCTTTCTTTCAATGATTC TAGCT	426
Db	18	HISLeutyValGlnMetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAla	37
QY	427	TATTCTTGCTGCTCTTTAAAATGGGGGGGATATGCAGCAGAAATCATGTCTCTCAAAGA	486
Db	38	ProCysPheAlaSerThr-----AlaPheThrValGluValIleWetProSerGlu	54
QY	487	ATTTACAGATGGGGAGACCTTAATCGTATATCTTCCCTATACTGTATAGGAGATCCGAGT	546
Db	55	AenPheAspGlySerSerGlyLysIle---PheProtyrThrThrLeuSerAspProArg	73
QY	547	GGGACTACTGTTTTTCTCGAGGAGAGTTAACATTAAAAAATCTTGACAATTTCTATTGCA	606
Db	74	GlyThrLeuCysIlePheSerSerGlyAspLeutyriIealaaSneUuaSpnaIalleSer	93
QY	607	GCTTTGGCCTTTAAAGTTGTTTTGGGAACCTATTATAGGAGAGTTTTACTGTTTTTAGGGAGAGA	666

Db 94 ArgThrSerSerSerCysPheSerAsnArgAlaGlyAlaLeuGluInLeuGlyLysGly 113
 Qy 667 CACTCGTTGACTTCGAGAACATACGACTTCTACAAATGGGGAGCTCTAAGTAATAGC 726
 Db 114 GlyValPheSerPheLeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerVal 133
 Qy 727 GCTGCTGATGCA-----CTGTTTACTATTGAGGGTTTTTAAAGAAATATCTCTTT 774
 Db 134 IleThrGlnAsnProGluLeuCysPheSerPheSerPheSerGlyPheSerGlnMetIlePhe 153
 Qy 775 TCCAATTGCAATTCATTCTTGGCCGTACTGCTGCAACGACTAAATAAGGGTAGCCAG 834
 Db 154 AspAsnCysGluSerLeuThr-----SerAsp 162
 Qy 835 ACTCGGAGCAACACTACACCGTCTAATGGTACTATTATTCTTAAACAGACTCTTTTG 894
 Db 163 ThrSerAlaSerAsnValIleProHisAlaSerAlaIleTyAlaThrThrProMetLeu 182
 Qy 895 TTACTCAATAATGAGAAGTCTCATTTCTATAGTAATTTAGTCTCTGGAGATGGGGAGCT 954
 Db 183 PheThrAsnAsnAspSerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAla 202
 Qy 955 ATAGATGCTAAGAGCTTAAAGCTTCAAGGAATTAGCAAGCTTTGTCTCTCCAAAGAAAT 1014
 Db 203 IleArgGlyThrSerIleThrIleGluAsnThrLysLysSerLeuLeuPheAsnGlyAsn 222
 Qy 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTTCTCTGCTATAGGCTAAC 1074
 Db 223 GlySerIleSerAsnGlyAlaLeuThrGlySerAlaAlaIleAsnLeuIleAsnAsn 242
 Qy 1075 GAGGCTCTATTGCTTTGAGCAAGTCTTCAGGAGTTCGAGGAGTAAAGGGAGGAGTCTGCT 1134
 Db 243 SerAlaProValIlePheSerThrAsnAlaThrGlyIleTyrGlyAlaIleTyrLeu 262
 Qy 1135 GTTCAGGATGGGACGACGAGGTGTCATCTACTTCAACAGAAAGTCCAGTAGTAGT 1194
 Db 263 Thr-----GlyGlySerMetLeuThrSer----- 270
 Qy 1195 TTTTCCAGAAATACGCGGTGAGTTTGTATGGAAACGTAGCCGAGTAGGAGGAGGATT 1254
 Db 271 --GlyAsnLeuSerGlyValLeuPheValAsnAsnSerSerArgSerGlyAlaAla 289
 Qy 1255 TACTCTACGGAAGCTTCTCTGTAATATGGAATAAAGCTTCTCAACAAATGTT 1314
 Db 290 TyrAlaAsnGlyAsnValThrPheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThr 309
 Qy 1315 GCTTCTCTCTGTT---TACATTGCTCTAAGCAACCAACAGTGGACAGGCTTCTAATACG 1371
 Db 310 AlaSerProGlnAsnSerLeuProAlaProThrProProProProProAlaValThr 329
 Qy 1372 AGTAATAATTACGGAGATGGAGGAGCTATCTTCTTAAGAAATGGTGCGCAAGCAGGATCC 1431
 Db 330 Pro----LeuLeuGlyTyrGlyGlyAlaIlePheCysThrProProAlaThrProProPro 348
 Qy 1432 AATAACTCTGATCAGTTTCTCTTGTATGGAGGGAGTAGTTTCTTTAGTAGCAATCTA 1491
 Db 349 ThrGlyVal---SerLeuThrIleSerGlyGluAsnSerValThrPheLeuGluAsnIle 367
 Qy 1492 GCTGCTGGGAAGGGAGCTATTATGCAAAAAGCTCTCGTTGTCTAACTGTGGCCCT 1551
 Db 368 AlaSerGluGlnGlyAlaLeuTyrGlyLysLysIleSerIleAspSerAsnLysSer 387
 Qy 1552 GTACAAATTTTAAAGAAATGCTAATGATGGTGAGCGATTATTTAGGAGAACTCTGGA 1611
 Db 388 ThrIlePheLeuGlyAsnThrAlaGlyLysGlyAlaIleAlaIleProGluSerGly 407
 Qy 1612 GAGCTCAGTTTCTGCTGATATCGAGATATTATTTTCATGGGAACTTTAAAGACA 1671
 Db 408 GluLeuSerLeuSerAlaAsnGlnGlyAspIleLeuPheAsnLysLeuSerIleThr 427
 Qy 1672 GCCAAAGAGAAATGCTGCCGATGTTAATGGCGTAAGTGTCTCTCAAGCCATTTGATG 1731
 Db 428 SerGly-----ThrProThrArgAsnSerIleHisPhe 438

Qy 1732 GGATCGGAGGGAATAACGACATTAAAGAGCTAAAGCAGCGCATCAGATTCTCTTTAAT 1791
 Db 439 GlyLysAspAlaLysPheAlaThrLeuGlyAlaThrGlnGlyTyrThrLeuTyrPheTyr 458
 Qy 1792 GATCCCATCGAGATGGCAACAGGAATAAACACGACGAGCTTCTCCAACTTCTAAAA 1851
 Db 459 AspProIle-----ThrSerAspAspLeuSerAlaAlaSerAlaAlaThrValVal 476
 Qy 1852 ATTAAC-----GATGGTGAAGGATACACACGGGATATTGTTTTT----- 1890
 Db 477 ValAsnProLysAlaSerAlaAspGly---AlaTyrSerGlyThrIleValPheSerGly 495
 Qy 1891 -----GCTAATCGAAGCAGTACTTCTTCAACAA 1917
 Db 496 GluThrLeuThrAlaThrGluAlaAlaThrProAlaAsnAlaThrSerThrLeuAsnGln 515
 Qy 1918 AATGTTACGATAGACGAGGAAGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAAT 1977
 Db 516 LysLeuGluLeuGluGlyThrLeuAlaLeuArgAsnGlyAlaThrLeuAsnValHis 535
 Qy 1978 TCTCTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGAGTACATGGGATTTT 2034
 Db 536 AsnPheThrGlnAspGluLysSerValIleMetAspAlaGlyThrThr-----Leu 553
 Qy 2035 GTAACTCCACCAACCCACCAAGCCCTCTCCGCTAATCAGTTGTATGATCAGCTTTTCCAAT 2094
 Db 554 AlaThrThrAsnGlyAlaAsnAsnThrAspGlyAla-----IleThrLeuAsnLys 570
 Qy 2095 CTGCATTGTCTCTCTTCTTCTTTGTAGCAACAAATGCAGTTACGAATCTCTTACCAAT 2154
 Db 571 LeuValIleAsnLeuAspSerLeuAspGlyThrLysAla----- 583
 Qy 2155 CCTCCAGCGCAAGATTCATCTCTCTGAGTCATTTGGT---AGCACAACTGCTGTTCTGTT 2211
 Db 584 -----AlaValValAsnValGlnSerThrAsnGlyAlaLeu 595
 Qy 2212 ACAATTAGTGGCTATCTTTTGGAGATTGGATGATACAGCTTATAGTAGTAGT 2271
 Db 596 ThrIleSerGlyThrLeuGlyLeuValLysAsnSerGlnAspCysCysAspAsnHisGly 615
 Qy 2272 TGGCTAGTCTTAAT---CAAAAATCAATGTCTGAAATTTACAGTTTAGGAGCTAAAGCC 2328
 Db 616 MetPheAsnLysAspLeuGlnValProIleLeuGluLeuLysAlaThrSerAsnThr 635
 Qy 2329 CCAGCTAATGCCCATCAGATTG---ACTTAGGGAATGAGATGCCCTAAGATATGGCTAT 2385
 Db 636 ValThrThrThrAspPheSerLeuGlyThrAsnGlyTyrGlnGlnSerProTyrGlyTyr 655
 Qy 2386 CAAGGAAGCTGGAAGCTTGGTGGATCCTAATAACGCAAAATATGCTCTTATCTGCTG 2445
 Db 656 GlnGlyThrTyrGluPheThrIleAspThrThr-----HisThrVal 670
 Qy 2446 AAAGCTACATGCACTAAAACTGGGTATAATCTCGGCTCAGCGAGTAGCTTCTTTGGTT 2505
 Db 671 ThrGlyAsnTyrLysLysThrGlyTyrLeuProHisProGluArgLeuAlaProLeuIle 690
 Qy 2506 CCAATAGTTTATGGGATCCATTTTAGATATACGATCGCGCATTCGCAATTCAGCAACGA 2565
 Db 691 ProAsnSerLeuTyrAlaAsnValIleAspLeuArgAlaValSerGlnAla-----Ser 708
 Qy 2566 AGTGTGATGGCCCTCTTAT---TGTCGAGGATTATGGTTTCTGGAGTTTCCGAATTC 2622
 Db 709 AlaAlaAspGlyLysGluAspValProGlyLysGlnLeuSerIleThrGlyIleThrAsnPhe 728
 Qy 2623 TTCTATCATGACCGCATGCTTTAGGTAGGATATCGGTATATAGTGGGGGTATTATCC 2682
 Db 729 PheHisAlaAsnHisThrGlyAspAlaArgSerTyrArgHisMetGlyGlyGlyTyrLeu 748
 Qy 2683 TTAGGAGCAACTCTTAC-----TTTGGATCATCGATGTTTGGTCTAGCATTTTACC 2733
 Db 749 Ile-----AsnThrTyrThrArgIleThrProAspAlaAlaLeuSerLeuGlyPheGly 766

polymorphic outer membrane protein g family - Chlamydothila pneumoniae (strain CWL029)
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: F72076
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: F72076
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <ARN>
A:Cross-references: UNIPROT:Q92896; UNIPARC:UPI000004707B; GB:AE001629; GB:AE001363; NID
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp_13
C:Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Alignment Scores: 5.63e-86 Length: 973
Pred. No.: 1365.50 Matches: 350
Score: 1365.50 Conservative: 171
Percent Similarity: 49.9%
Best Local Similarity: 33.5% Mismatches: 420
Query Match: 17.3% Indels: 103
DB: 27 Gaps: 27

US-10-701-844-1 (1-4435) x F72076 (1-973)

QY	382	ATGCAACGCTTTCCATAGTCTTTCTTTCAATGATCTAGCTTATCTGCTCTCT	441
DB	1	MetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaProCysPheAlaSer	20
QY	442	TTAAATGGGGGGGATATGACAGAGAAATCATGCTCTCTCAAGAAATTCACGATGGGAG	501
DB	21	Thr-----AlaPheThrValGluValIleMetProSerGluAsnPheAspGlySer	37
QY	502	ACGTTAACTGATCATCTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTGTTT	561
DB	38	SerGlyLysIle---PheProTyrThrThrLeuSerAspProArgGlyThrLeuCysIle	56
QY	562	TCGCGAGAGATTAACATTAATAAATCTTGCAATCTTATGCAAGCTTTGCTTAAAGT	621
DB	57	PheSerGlyAspLeuTyrIleAlaAsnLeuAspAsnAlaIleSerArgThrSerSer	76
QY	622	TGTTTGGGAACCTATTAGGAGTTTACCTGTTTATAGGAGAGACACTCGTTGACTTTC	681
DB	77	CysPheSerAsnArgAlaGlyAlaLeuGlnIleLeuGlyLysGlyValPheSerPhe	96
QY	682	GAGACATACGACTCTACAAATGGGCGAGCTTAAGTAATAGCGCTGCTGATGGA---	738
DB	97	LeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIleThrGlnAsnPro	116
QY	739	-----CTGTTTACTATGAGGGTTTAAAGAAATTTATCTTTTCCAAATTCGAATTC	789
DB	117	GluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAspAsnCysGluSer	136
QY	790	TTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGGTACCGACACTCCGACGACA	849
DB	137	LeuThr-----SerAspThrSerAlaSerAsn	145
QY	850	TCACACGCTCTAATGCTATTATTCTAARACAGATCTTTTGTCTACTCAATAATGAG	909
DB	146	ValIleProHisAlaSerAlaIleTyrAlaThrThrProMetLeuPheThrAsnAsnAsp	165
QY	910	AAGTTCTCATCTAGTAATTTAGTCTCGAGATGGGGAGCTATAGATGCTTAAGAGC	969
DB	166	SerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAlaIleArgGlyThrSer	185
QY	970	TTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAATATCTGCTCAAGCTGAT	1029
DB	186	IleThrIleGluAsnThrLysLysSerLeuLeuPheAsnGlyAsnGlySerIleSerAsn	205
QY	1030	GGGGAGCTTGTCAAGTAGTCACCAAGTTCCTCTGCTATGGCTAACGAGGCTCTATTGCC	1089

DB	206	GlyGlyAlaLeuThrGlySerAlaAlaIleAsnLeuIleAsnAsnSerAlaProValIle	225
QY	1090	TTTGTAGCGAATGTTGCAGGAGTAAGGGGAGGATTGCTGCTGTTCCAGGATGGCGAG	1149
DB	226	PheSerThrAsnAlaThrGlyIleTyrGlyAlaIleTyrLeuThr-----	241
QY	1150	CAGGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACT	1209
DB	242	---GlyGlySerMetLeuThrSer-----GlyAsnLeuSer	252
QY	1210	GGGTAGAGTTGATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAC	1269
DB	253	GlyValLeuPheValAsnAsnSerSerArgSerGlyAlaIleTyrAlaAsnGlyAsn	272
QY	1270	GTTGCTTCTCGAATAATGGAACCTGTTTCTCAACATGTTGCTTCTCTCTCT	1326
DB	273	ValThrPheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThrAlaSerProGlnAsn	292
QY	1327	TACATTGCTGCTAAGCAACCAACAGTGGACAGGCTTCTAATACAGTAATTAATTCGGA	1386
DB	293	SerLeuProAlaProThrProProProProAlaValThrPro---LeuLeuGly	311
QY	1387	GATGAGAGGCTATCTCTGTAAGAAATGTTGGCGCAGCGAGTCCAACTCTGGATCA	1446
DB	312	TyrGlyGlyAlaIlePheCysThrProProAlaThrProProThrGlyVal---Ser	330
QY	1447	GTTTCTTGTAGGAGGAGGAGTGTCTTTCTTAGCAATGATGCTGCGGAAGGG	1506
DB	331	LeuThrIleSerGlyGluAsnSerValThrPheLeuGluAsnIleAlaSerGluGlnGly	350
QY	1507	CGAGCTATTATGCAAAAGCTCTCGTGTCTAACTGTGGCCCTGTACAAATTTTAAAG	1566
DB	351	GlyAlaLeuTyrGlyLysIleSerIleAspSerAsnLysSerThrIlePheLeuGly	370
QY	1567	ATATGCTTAATGATGCTGGAGCGGATTTATTTAGGAGATCTCGAGAGCTCAGTTTATCT	1626
DB	371	AsnThrAlaGlyLysGlyAlaIleAlaIleProGluSerGlyGluLeuSerLeuSer	390
QY	1627	GCTGATTATGAGATTTATTTTCATGCGGAATCTTAAAGAACACAGCCAAAGAGATGCT	1686
DB	391	AlaAsnGlnGlyAspIleLeuPheAsnLysAsnLeuSerIleThrSerGly-----	407
QY	1687	GCGGATTTAATGGCGTAACTGTCTCTCACAGCCATTTTCGATGGGATCGGAGGAAA	1746
DB	408	-----ThrProThrArgAsnSerIleHisPheGlyLysAspAlaLys	421
QY	1747	ATAACGACATTAAGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCATCGAGATG	1806
DB	422	PheAlaThrLeuGlyAlaThrGlnGlyTyrThrLeuTyrPheTyrAspProIle-----	439
QY	1807	GMAACGGAATAACACCGCCAGCGAGTCTTCCAACTTCTAAATAATTAAC-----	1857
DB	440	ThrSerAspAspLeuSerAlaAlaSerAlaAlaThrValValAsnProLysAla	459
QY	1858	-----GATGGTGAAGTACACAGGGGATATTGTTTTT-----	1890
DB	460	SerAlaAspGly---AlaTyrSerGlyThrIleValPheSerGlyGluThrLeuThrAla	478
QY	1891	-----GCTAATGGAACGAGTCTTTGTACCAAAATGTTACGATAGAG	1932
DB	479	ThrGluAlaAlaThrProAlaAsnAlaThrSerThrLeuAsnGlnLysLeuGluLeuGlu	498
QY	1933	CAAGGAGGATGTTCTTCTGTAAGGCAAAATTTATCAGTGAATTTCTCTAAGTCAAGACA	1992
DB	499	GlyGlyThrLeuAlaLeuArgAsnGlyAlaThrLeuAsnValHisAsnPheThrGlnAsp	518
QY	1993	GGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCCAACCA	2049
DB	519	GluLysSerValIleMetAspAlaGlyThr-----LeuAlaThrThrAsnGly	536
QY	2050	CCAAACAGCTCTCTGCGCTAATCAATGATGATCAGCTTTCCCAATCTGCAATTTGCTCTT	2109

```
Db 537 AlaAenAenThrAspGlyAla-----lleThrLeuAenLysLeuValleAenLeu 553
Qy 2110 TCTTCTTTGTAGCAAAACAATGCAATAGCAATCTCTTACCAATCTCCAGCGCAAGAT 2169
Db 554 AspSerLeuAspGlyThrLysAla-----
Qy 2170 TCTCATCTCGAGTCATTGGT---AGCACAACTGCTGCTTCTGTGTACAAATAGTGGGCT 2226
Db 562 -----AlaValAenValGlnSerThrAenGlyAlaLeuThrileSerGlyThr 578
Qy 2227 ATCTTTTGTAGGATTTCGATGATACAGCTTATGATAGTATGATGATTGGCTAGGTTCTAAT 2286
Db 579 LeuGlyLeuValLysAenSerGlnAspCysCysAspAenHisGlyMetPheAenLysAsp 598
Qy 2287 ---CAAAAAATCAATGCTGAAATACAGTTAGGCAATAAGCCGCCAGCTAATGCCCA 2343
Db 599 LeuGlnValProileLeuLeuLysAlaThrSerAenThrValThrThrAsp 618
Qy 2344 TCAGATTGG---ACTCTAGGCAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAG 2400
Db 619 PheSerLeuGlyThrAenGlyTyrGlnSerProfyGlyTyrGlnGlyThrTrpGlu 638
Qy 2401 CTTCGCTGGGATCTTAATACAGCAAAATAATGCTCCTTATCTCTGAAAGCTACATGGACT 2460
Db 639 PheThrileAspThrThr-----HisThrValThrGlyAenTrpLys 653
Qy 2461 AAAACTGGGTAAATCTGGGCCCTGAGCGAGTAGTCTCTTTGGTTCCAAATAGTTATGG 2520
Db 654 LysThrGlyTyrLeuProHisProGluArgLeuAlaProLeuileProAenSerLeuTrp 673
Qy 2521 CGATCCATTTAGATATAGTCTCGCATTCAGCAATTCAGCAAGTGTGGTGGCGC 2580
Db 674 AlaAenValileAspLeuArgAlaValSerGlnAla-----SerAlaAlaAspGlyGlu 691
Qy 2581 TCTTAT---TGTCCGAGGATTATGGGTTCTGGAGTTTCGAAATTTCTTCTATCATGACCGC 2637
Db 692 AspValProGlyLysGlnLeuSerileThrGlyleThrAenPheHisAlaAenHis 711
Qy 2638 GATGCTTTAGGTCAGGATATCGGTATATAGTGGGGTTATTCCTTAGGAGCAAACTCC 2697
Db 712 ThrGlyAspAlaArgSerTyrArgHisMetGlyGlyTyrLeuile-----AenThr 729
Qy 2698 TAC-----TTTGGATCATCGATGTTGTGCTAGCATTTACCGAAGTATTTGGTAGA 2748
Db 730 TyrThrArgileThrProAspAlaAlaLeuSerLeuGlyPheGlyGlnLeuPheThrLys 749
Qy 2749 TCTAAGATTATGATGTGTCGTTCCCAATCATCATGCTTGCATAGGATCGTTTATCTA 2808
Db 750 SerLysAspTyrLeuValGlyHisGlyHisSerAenValTyrPheAlaThrValTyrSer 769
Qy 2809 TCTACCCCAACAGCTTTATGTGGA---TCCTATTTTGTTCGGAGATGCGTTTATCCGTGCT 2865
Db 770 AenileThrLysSerLeuPheGlySerSerArgPhePheSerGlyThrSerArgVal 789
Qy 2866 AGCTACGGGTTTGGGAATCAGCATATGAACCTCATATACATATTCGAGAGAGCGCAT 2925
Db 790 ThrTyrSerArgSerAenGluLysValLysThrSerTyrThrLysLeuLeuProLysGlyArg 809
Qy 2926 GTTCGTTGGGATAAATCTGCTGCTGGAGATTTGGAGCGGATTAACCGATTTCGTGATT 2985
Db 810 CysSerTrpSerAenAenCysTrpLeuGlyLysLeuGlyAenLeuProleThrLeu 829
Qy 2986 ACTCCATCTAAGCTCTATTGTAATGAGTTGCGTCTCTTCGTCGAAGCTGAGTTTCTTAT 3045
Db 830 SerSerArgileLeuAenLysGlnleileProPheValLysAlaGluValAlaTyr 849
Qy 3046 GCCGATCATGAATCTTTACAGAGAGCGCATCAAGCTCGGCATTTCAGAGCGGCAT 3105
Db 850 AlaThrHisGlyGlyleGlnGlnAenThrProGluGlyArgilePheGlyHisGlyHis 869
Qy 3106 CTCCTAAATCTATAGTCTCTTCGAGTGAAGTTTGTATCGATGTTCTAGTACACATCCT 3165
Db 870 LeuLeuAenValAlaValProValGlyValArgPheGlyLysAenSerHisAenArgPro 889
```

```
Qy 3166 AATAAATATAGCTTTATGGCGGCTTATATCTGTGATGCTTATCCACCAATCTCTGGTACT 3225
Db 890 AspPheTyrThrileileValAlaTyrAlaProAspValTyrArgHisAenProAspCys 909
Qy 3226 GAGACAACGCTCCATCCATCAAGAGACATGGACAACAGATGCTTTTCATTTAGCAAGA 3285
Db 910 AspThrThrLeuProleAenGlyAlaThrTrpThrSerileGlyAenAenLeuThrArg 929
Qy 3286 CATGAGCTTGTGGTTAGAGGATCTATGTATGCTTCTCTAAACAAGTAATAAGAGTATAT 3345
Db 930 SerThrLeuLeuValGlnAlaSerSerHisThrSerValAenAenValLeuGluilePhe 949
Qy 3346 GGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGGTGCAGGAGT 3405
Db 950 GlyHisCysGlyCysAspIleleArgArgThrSerArgGlnTyrThrLeuAspIleGlySer 969
Qy 3406 AGACTCCGGTTC 3417
Db 970 LysLeuArgPhe 973

RESULT 7
A:811723
Polymorphic membrane protein H family TC0264 [imported] - Chlamydia muridarum (strain N1)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Dec-2002
C:Accession: A811723
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684955
A:Accession: A811723
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-983 <TST>
A:Cross-references: UNIPARC:UPI0000057868; GB:AE002293; GB:AE002160; NID:G7190298; PIDN:
A:Experimental source: strain N199 (MoPn)
C:Genetics:
A:Gene: TC0264
C:Superfamily: Polymorphic membrane protein H family

Alignment Scores:
Pred. NO.: 3.95e-73 Length: 983
Score: 1179.00 Matches: 232
Percent Similarity: 78.0% Conservative: 27
Best Local Similarity: 69.9% Mismatches: 56
Query Match: 15.0% Indels: 17
DB: 2 Gaps: 3

US-10-701-844-1 (1-4435) x A811723 (1-983)
Qy 3442 GTGTAGCGATGCTCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTAGCTTGTGTGT 3501
Db 1 MetLeuValMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 20
Qy 3502 TCCTATTGTTATGATTGCGAGCTCTCTCAAGTGTAAAGCTTAATGTAAACCACTCCT 3561
Db 21 SerTyrSerTyrGlyLeuAlaSerSerProGlnValLeuThrProAenValilelePro 40
Qy 3562 TTTAAGGAGACGATGTTTACTTGAATGGAGACTGCGCTTTTGTCAATGCTTATGACGGA 3621
Db 41 PheLysGlyAspAspIleTyrLeuAenGlyAspCysValPheAlaSerileTyrAlaGly 60
Qy 3622 GCTCAAGAAGGTTTCGATTATCTCAGCTAAATGGCGACAATTTAAACGATTACCGGACAAAAC 3681
Db 61 AlaGluGlnGlySerileleSerAlaAenGlyGlnAenLeuThrileValGlyGlnAen 80
Qy 3682 CATATTTATCATTTACAGATTCTCAAGGGCCAGTCTCTCAAAATATTATGCTTCAATTTCA 3741
Db 81 HisThrLeuSerPheThrAspSerGlnGlyProAlaLeuGlnAenCysAlaPheileSer 100
Qy 3742 GCAGGAGACACTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAAATGTTTCT 3801
```

101 AlaGluGluLysIleSerLeuArgAspPheSerSerLeuLeuPheSerLysAsnValSer 120
3802 TGGCGGAAGAAAGGAATGATCTCCGGGAAACCGTCAGTATTTCCGGAGCAGCGGAAGTG 3861
121 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyLysAspSerIle 140
3862 ATTTTCTGGGATAACTCCGTCGGGTATTTCTCTTTATCTACTGTGCAACCTCATCATCA 3921
141 ValPheLysAspAsnSerValGlyTyrSerSerLeu-----ProSerValGlyGln 157
3922 ACTCCGCTGCTCCACAGTATGATGTCGCGAAAGGTCATATTTTCTGTAGAGACT 3981
158 ThrProThrThrProTleValGlyAspValLeuLysGlySerIlePheCysValGluThr 177
3982 AGTTTCGAGATCTCAGCGCTCAAAAAGGGTGCATGTTTCGATAATAATGTCGCGGAATTTTC 4041
178 GlyLeuGluIleSerGlyValLysLysGluLeuValPheAspAsnThrAlaGlyAsnPhe 197
4042 GGAACAGTTTTCGAGTAAAGATAATAATGCTGCTGGTGGAGCGAGTGGGTTCCGC 4101
198 GlyAlaValPheCysSer-Arg-----A 205
4102 TACACCATCAAGTAGCATTTTACAGTTAAAACTGTAAAGGAAAGTTTCTTTCCACAGA 4161
205 AlaGlnGlyAspThrThrPheThrValLysAspCysLysGlyLysIleLeuPheGlnAs 225
4162 TAACGTAGCTCTTCGGAGCGCGGTGTTTATAAGGCAATGTCCTTTTCAAGACAA 4221
225 AsnValGlySerCysGlyGlyValIleTyrLysGlyGluValLeuPheGlnAspAs 245
4222 TGAAGGAGGATATCTTCCGAGGGAACACAGCATACGATGATTAAGGATTTCTGTGTC 4281
245 nGluGlyGluMetLeuPheArgGlyAsnSerAlaHisAspAspLeuGlyIleLeuAspAl 265
4282 TACTAATACGATCAGAAATACGAGAGACAGAGCGCGTGGAGAGTATTTTGTCTCCAGA 4341
265 a--AsnProGlnProProThrGluValGlyGlyGlyValIleCysThrProGln 284
4342 TGATTCGTAAAGTTTCAAGGCAATAAGGTTCTATTGTTTGTATTACAACTTTGCAAA 4401
284 uLysThrValThrPheLysGlyAsnLysGlyProIleThrPheAspTyrAsnPheAlaLys 304
4402 AGGCAGAGCGGAGCATCTCTAACGAAGAATTC 4435
304 sGlyArgGlyGlyAlaIleGlnSerGlnThrPhe 315

RESULT 8
G86546
polymorphic outer membrane protein G family [imported] - Chlamydomonas pneumoniae (strain
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise,
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: G86546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <STO>
A:Cross-references: UNIPROT:Q9R865; UNIPARC:UPI000002PFF0; GB:BA000008; NID:g8978821; I
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_10
C:Superfamily: Chlamydomonas pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 9, 17e-68 Length: 928
Score: 1101.00 Matches: 324
Percent Similarity: 44.5% Conservative: 150
Best Local Similarity: 30.5% Mismatches: 402
Query Match: 14.0% Indels: 188

RESULT 9

G81591
polymorphic membrane protein G family CP0303 [imported] - Chlamydomophila pneumoniae (strain G81591)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: G81591
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: G81591
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <REA>
A:Cross-references: UNIPROT:Q9RB65; UNIPARC:UPI000002FFP0; GB:AE002161; NID
A:Experimental source: strain AR39, HL cells
C:Genetic:
C:Gene: CP0303
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:

Pred. No.:	9.17e-68	Length:	928
Score:	1101.00	Matches:	324
Percent Similarity:	44.5%	Conservative:	150
Best Local Similarity:	30.5%	Mismatches:	402
Query Match:	14.0%	Indels:	188
DB:	2	Gaps:	34

US-10-701-844-1 (1-4435) x G81591 (1-928)

QY	382	ATGCAACGCTTTCCATAGATCTCTTTTCAATGATCTAGCT---TATCTTCTGCTGC	438
DB	1	MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys	20
QY	439	TCCTTAATGGGGGGATATGCGAGCAATCATGTTCTCTCAAGAAATTTACGATGGG	498
DB	21	SerThrValPheAlaAla---ThrAlaGluAsnIleGlyProSerAspSerPheAspGly	39
QY	499	GAGACGTTACTCTATCATTTCCCTATCTACTGTTATAGGAGATCCGAGTGGGACTACTGTT	558
DB	40	SerThrAsnThrGlyThr-----TyrThrProLysAsnThrThrThrGlyIleAspTyr	57
QY	559	TTTTCTGCAGAGATTAACTAAATAAATCTTGACAAATCTATGCGCTTTGCTTCA	618
DB	58	ThrLeuThrGlyAspIleThrLeuGlnAsnLeuGlyAspSer---AlaAlaLeuThrLys	76
QY	619	AGTTGTTTGGGAACCTATTAGGGAGTTTACTGTTTTCAGGAGAGACACTGTTGACT	678
DB	77	GlyCysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSer	96
QY	679	TTGAGAACATACGACTCTTACAAATGGGCGAGCTCTAAGTAATACGCTGCTGATGGA	738
DB	97	PheLeuAsnIleLysSerSerAlaGluGlyAlaAlaLeuSer---ValThrAspLys	115
QY	739	CTGTTTACTATTAGGGTTTAAAGAAATATCTTTTCCAAATTCATTTACTTTGCC	798
DB	116	AsnLeuSerLeuThrGlyPheSerSerLeuThrPhe-----LeuAlaAla	130
QY	799	GTACTGCTGTCGACGACTAATAGGGTAGCCGAGACTCCGACGACCAACATCTACACCG	858
DB	131	ProSerSerValIleThrThrProSerGly-----	140
QY	859	TCTAATGCTACTATTATTCTTAAACAGATCTTTTGTACTCAATAATGAGAGTTCTCA	918
DB	141	---LysGlyAlaValLysCysGlyGlyAspLeuThrPheAspAsnAsnGlyThrIleLeu	159
QY	919	TTCTATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGTAAGAGCTTAAACGGTT	978
DB	160	PheLysGlnAspTyrCysGluGluAsnGlyGlyAlaIleSerThrLysAsnLeuSerLeu	179
QY	979	CAAGGAATTAGCAAGCTTTTGTTGTTCTTCCAGAAATATACTGCTCAAGCTGATGGGGAGCT	1038

DB	180	LysAsnSerThrGlySerIleSerPheGluGlyAsnLysSerSerAla-----	195
QY	1039	TGTCAGTAGTACACGAGTTTCTCTGCTATGCTTAACGAGGCTCTTATGCTTGTAGCG	1098
DB	195	-----	195
QY	1099	AATGTTTCAGGAGTAAGAGGGGAGGATGCTGCTTTCAGGATGGCAGCAGGAGTG	1158
DB	196	-----ThrGlyLysGlyAlaIleCysAla-----	205
QY	1159	TCATCATCTACTTCAACAGAGATCCAGTAGTATGTTTTCAGAAATACTCGC---GTA	1215
DB	206	-----ThrGlyThr-----ValAspIleThrAsnAsnThrAlaProThr	218
QY	1216	NGTTTGTATGGAACGTAGCCGA---GTAGGAGGAGGATTTACTCTTACGGGAACGTT	1272
DB	219	LeuPheSerAsnAsnIleAlaGluAlaAlaGlyGlyAlaIleAsnSerThrGlyAsnCys	238
QY	1273	GCTTTCTCGAATAATGGAACCTGTTTCTTCAACAATGTTGCTTCTCTCTTTTACATT	1332
DB	239	ThrIleThrGlyAsnThrSerLeuValPheSerGluAsn-----	251
QY	1333	GCTGTAAGCAACCAAGTGGACAGGCTTCTAATACGAGTAATAATTACGAGATGGA	1392
DB	252	-----SerValThrAlaThrAlaGlyAsnGly	260
QY	1393	GGAGCTATCTTCTGTAAGAATGGTGGCAAGCAGGATCCAAATACTCTGATCAGTTTCC	1452
DB	261	GlyAlaLeu-----SerGlyAspAlaAspValThr	270
QY	1453	TTTGATGAGAGGAGTAGTCTTTTCTTAGTACAAATGCTGCTGGGAAAGGGAGCT	1512
DB	271	IleSerGlyAsnGlnSerValThrPheSerGlyAsnGlnAlaValAlaAsnGlyGlyAla	290
QY	1513	ATTATGCAAAAGCTCTCGGTTGCTAAC-----TGTGGCCCTGTACAAATTTTAAGG	1566
DB	291	IleTyrAlaLysLysLeuThrLeuAlaSerGlyGlyGlyIleSerPheSerAsn	310
QY	1567	ATATTC-----CCTAATGATGCTGGAGCGATTTATTAGGAGAACTCTGGA	1611
DB	311	AsnIleValGlnGlyThrThrAlaGlyAsnGlyGlyAlaIleSerIleLeuAlaAlaGly	330
QY	1612	GAGCTCAGTTTATCTGCTGATTTATGAGATATTTTTCGATGGGAAT-----	1659
DB	331	GluCysSerLeuSerAlaGluAlaGlyAspIleThrPheAsnGlyAsnAlaIleValAla	350
QY	1660	---CTTAAAGAACACGCCAAAGAGATGCTGCGGATGTTTAAATGGCGTAACTGTGCTCA	1716
DB	351	ThrThrProGlnThrThrLysArgAsnSerIleAspIle-----	363
QY	1717	CAAGCATTTCCGATGGGATCGGGAGGAAATAACGACATTAAGACCTAAGCAGGCGCAT	1776
DB	364	-----GlySerThrAlaLysIleThrAsnLeuArgAlaIleSerGlyLys	378
QY	1777	CAGATTCTCTTAAATCATCCATCGAGATGCTGCGGATGTTTAAATGGCGTAACTGTCT	1836
DB	379	SerIlePhePheThrAspProIleThrAla-----AsnThrAlaAlaAspSer	394
QY	1837	TCCAAATCTCTAAAAATTAACGATGCTGAAGGA-----TACACAGGGGAT	1881
DB	395	ThrAspThrLeuAsnLeuAsnLysAlaAspAlaGlyAsnSerThrAspTyrSerGlySer	414
QY	1882	ATGTTTCTGCT-----AATGNAAGC	1902
DB	415	IleValPheSerGlyGlyLysLeuSerGluAspGluAlaLysValAlaAspAsnLeuThr	434
QY	1903	AGTACTTGTACCAATTTGATAGACAGGAGGATGCTTCTCTCTGTAAGAGGCA	1962
DB	435	SerThrLeuLysGlnProValThrLeuThrAlaGlyAsnLeuValLeuLysArgGlyVal	454
QY	1963	AAATTATCAGTGAATTTCTTAAGTCAAGCA---GGTGGGAGTCTGTATATGGAAGCTGGG	2019


```
Db 455 ThrLeuaspThrLysGlyPheThrGlnThrAlaGlySerSerValIleMetAspAlaGly 474
Qy 2020 AGTACATGGGATTTTGTAACTCCACACCACACACAGCCCTCTCGCGCTAAATCAGTTG 2079
Db 475 ThrThrLeu-----LysAlaSerThrGluGlu 483
Qy 2080 ATCAGCTTTCATCTGCATTTGTCTCTTCTCTCTTTTGTAGCAAAATGCGAGTTACG 2139
Db 484 ValThrLeuThrGlyLeuSerIleProValAspSerLeu----- 496
Qy 2140 AATCTCTACCAATCTCCAGCGCAAGATTCTCATCTCGCAGTCATTTGGTAGCACAAC 2199
Db 497 -----GlyGluGlyLysLysValIleAlaAlaSerAlaAla 509
Qy 2200 GCTGGTCTCTGTTACAATTAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTAT 2259
Db 510 SerLysAenValAlaLeuSerGlyProIleLeuLeuLeuAenGlnGlyAenAlaTyr 529
Qy 2260 GATAGGTATGATGGCTAGGTCTTAATCAAAAAATCAATGCTCCTGAATTACAG---TTA 2316
Db 530 GluAenHisAsp---LeuGlyLysThrGlnAspPheSerPheValGlnLeuSerAlaLeu 548
Qy 2317 GGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCGCTAAG 2376
Db 549 GlyThrAlaThrThrAspValProAlaValProThrValAlaThrPro---ThrHis 567
Qy 2377 TATGGCTATCAAGGAGCTGGAAGCTGCGGGATCCTCAATACAGCAAAATAGTGTCT 2436
Db 568 TyrGlyTyrGlnGlyThrTyrGlyMetThrTyrValAspAspThrAlaSerThr---Pro 586
Qy 2437 TATACTCTGAAGCTACA-----TGGACTAAACTGGGTATATATCTCGGCGCTGAGGCA 2490
Db 587 LysThrLysThrAlaThrLeuAlaTyrThrAenThrGlyTyrLeuProAenProGluArg 606
Qy 2491 GTAGCTCTTGGTTCCTCAATAGTTATGGGATCCATTTAGATATACGATCTCGCAT 2550
Db 607 GlnGlyProLeuValProAenSerLeuTyrGlySerPheSerAspIleGlnAlaIleGln 626
Qy 2551 TCAGCAATTCAGCAAGTGTGTGATGGCGCTCTTATTCGAGGATATATGGTTTCTGGA 2610
Db 627 GlyValIleGluArgSerAlaLeuThrLeuLeuCysSerAspArgGlyPheTyrAlaAlaGly 646
Qy 2611 GTTTCGAATTTCTTATCATGACCGCGATGCTTTAGGTACGGGATATCGGTATATTAGT 2670
Db 647 ValAlaAenPheLeuAspLysAspLysLysGlyGluLysArgLysTyrArgHisLysSer 666
Qy 2671 GGGGGTATTCTCTTA---GGAGCAACTCTTACTTTGGATCATCGATGTTGGTCTAGCA 2727
Db 667 GlyGlyTyrAlaIleGlyGlyAlaAlaGlnThrCysSerGluAenLeuIleSerPheAla 686
Qy 2728 TTTTACCGAATATTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAAATCATCATGCT 2787
Db 687 PheCysGlnLeuPheGlySerAspLysAspPheLeuValAlaLysAenHisThrAspThr 706
Qy 2788 TGCATAGGATCGGTTATCTATCT-----ACC 2814
Db 707 TyrAlaGlyAlaPheTyrIleGlnHisIleThrGluCysSerGlyPheIleGlyCysLeu 726
Qy 2815 CAACAGCTTTATGCGATCTTATTTGTCGGAGATGCGTTTATC-----CGTGTAGC 2868
Db 727 LeuAspLysLeuProGlySerTyrSerHisLysProLeuValLeuGluGlyGlnLeuAla 746
Qy 2869 TACGGGTTTGGGAATCAGCATATGAACCTCATATACATTTTGCAGAGGAGCGCATGTT 2928
Db 747 TyrSerHisValSerAenAspLeuLysThrLysTyrThrAlaTyrProGluValLysGly 766
Qy 2929 COTTCGGATATAACTGCTGCTGGAGATGAGCGGGATACCGATGTTGATTACT 2988
Db 767 SerTyrGlyAenAenAlaPheAenMetMetLeuGlyAlaSerSerHisSerTyr----- 784
Qy 2989 CCATCTAGCTCTAT---TTGAATGAGTTGCGTCTCTTCGTCAAGCTGAGTTTCTTAT 3045
Db 785 ProGluTyrLeuHisCysPheAspThrTyrAlaProTyrIleLysLeuAenLeuThrTyr 804
```

```
Qy 3046 GCCCATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGGCATTTCAAGCGGACAT 3105
Db 805 IleArgGlnAspSerPheSerGluLysGlyThrGluGlyArgSerPheAspAspSerAen 824
Qy 3106 CTCCTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCT 3165
Db 825 LeuPheAenLeuSerLeuProIleGlyValLysPheGluLysPheSerAspCysAenAsp 844
Qy 3166 AATAAATATAGCTTTATGGCGGCTTATCTGTGATGCTTATCGCACCATCTCTCTGTACT 3225
Db 845 PheSerTyrAspLeuThrLeuSerTyrValProAspLeuIleArgAenAspProLysCys 864
Qy 3226 GAGACAACGCTCTCATCCATCAAGACATGAGCAACAGATGCTTTTCAATTAGCAAGA 3285
Db 865 ThrThrAlaLeuValIleSerGlyAlaSerTyrProGluThrTyrAlaAenAenLeuAlaArg 884
Qy 3286 CATGAGCTTCTGGTATAGGATCTATCTATGCTCTCTCAACAGTAATATAGAAATATAT 3345
Db 885 GlnAlaLeuGlnValArgAlaGlySerHisTyrAlaPheSerProMetPheGluValLeu 904
Qy 3346 GGCCATCGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTTCGAGGAGT 3405
Db 905 GlyGlnPheValPheGluValArgGlySerSerArgIleTyrAenValAspLeuGlyGly 924
Qy 3406 AGATCCCGGTTTC 3417
Db 925 LysPheGlnPhe 928

RESULT 10
F81591
polymorphic membrane protein G family CP0302 [imported] - Chlamydomphila pneumoniae (stra
C;Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
C;Accession: F81591
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: AB1500; MUID:20150255; PMID:10684935
A;Accession: F81591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-949 <R>A>
A;Cross-references: UNIPARC:UPI00001655FA; GB:AB002192; GB:AB002161; NID:g7189226; PIDN:
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0302
C;Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 3.84e-67 Length: 949
Score: 1092.00 Matches: 316
Percent Similarity: 42.7% Conservative: 154
Best Local Similarity: 28.7% Mismatches: 409
Query Match: 13.9% Indels: 222
DB: 2 Gaps: 31

US-10-701-844-1 (1-4435) x F81591 (1-949)
Qy 322 ATTAATATTATTAATTTTATGAAGCGGAGTAATTAATTTATCTCTCAGCTTTGTGTG 381
Db 2 IleTyrLeuPheCysPheTyrIleAspAlaAenSerSerLeuLysAenLysSerIleThr 21
Qy 382 ATGCAAGCGTCTTCCATAGCTTCTTCTTCAATGATTTCTAGCTTATCTTCTTTCG 435
Db 22 MetLysThrSerIleProTyrValLeuValSerSerValLeuAlaPheSerCysHisLeu 41
Qy 436 TGCTCTTTTAAATGGGGGGGATATGACGACAAATCATGTTCTCAAGGAATTTACGAT 495
Db 42 GlnSerLeuAlaAen-----GluGluLeuLeuSerProAspAspSerPheAen 57
Qy 496 GGGGAGGAGCTTAATCTGATCATTT-----CCCTATCTGTTATA 534
```

```
Db      58 GlyAsnIleAspSerGlyThrPheThrProLysThrSerAlaThrThrThrThrSerLeuThr 77      |||      : : :      |||      : : : : :
Qy      535 GGAGATCCGAGTGGGACTACTGTTTTCCTGCGAGGAGATTAAACATTAATAATCTTGAC 594      |||      : : :      |||      : : : : :
Db      78 GlyAsp-----ValPhePheThrGluProGlyLysGlyThr----- 89      |||      : : :      |||      : : : : :
Qy      595 AATTCATTATGCAGCTTTCCTTA-----AGTTGTTTGGGAACCTATTAGGAGCTTTT 648      |||      : : :      |||      : : : : :
Db      90 -----ProLysSerAspSerCysPheLysGlnThrThrAspAsnLeu 103      |||      : : :      |||      : : : : :
Qy      649 ACTGTTTATAGGAGAGGACACTCGTTGACTTTCGAGACATACGGACTTCTCAAAATGGG 708      |||      : : :      |||      : : : : :
Db      104 ThrPheLeuGlyAsnGlyHisSerLeuThrPheGlyPheIleAspAlaGlyThrHisAla 123      |||      : : :      |||      : : : : :
Qy      709 CGAGCTTAAGTAATAGCGTCTGATGAGCTGTTTACTATTGAGGGTTTAAAGAAATTA 768      |||      : : :      |||      : : : : :
Db      124 GlyAlaAlaAlaSerThrAlaAsnLysAsnLeuThrPheSerGlyPheSerLeuLeu 143      |||      : : :      |||      : : : : :
Qy      769 TCCTTTTCCAAATTCATTACTTTCCTGCTACTGCTGCTCAACGACTAATAAGGCT 828      |||      : : :      |||      : : : : :
Db      144 SerPheAspSer----- 147      |||      : : :      |||      : : : : :
Qy      829 AGCCAGACTCCGACGACCAACATCTACACCGTCTAATGGTACTATTTATTCTAAACAGAT 888      : : : : :      |||      : : :      |||      : : : : :
Db      148 -----SerProSerThrThrValThrThrGlyGlnGlyThrLeuSerSerAlaGlyGly 165      : : : : :      |||      : : :      |||      : : : : :
Qy      889 CTTTGTCTACTCAATAATGAGAAGTCTCATCTTCTATAGTAATTTAGTCTCTGAGATGGG 948      : : :      |||      : : :      |||      : : : : :
Db      166 ValAsnLeuGluAsnIleArgLysLeuValAlaGlyAsnPheSerThrAlaAspGly 185      : : :      |||      : : :      |||      : : : : :
Qy      949 GAGCTATAGACTTAAGACTTAACGGTTCAGGAATTAGCAAGCTTTGTGCTCTCCAA 1008      : : :      |||      : : :      |||      : : : : :
Db      186 GlyAlaIleLysGlyAlaSerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSer 205      : : :      |||      : : :      |||      : : : : :
Qy      1009 GAAATATGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCCACGATTCCTCTGCTATG 1068      |||      : : :      |||      : : : : :
Db      206 AsnAsnSerSerThrLysGlyAla----- 215      |||      : : :      |||      : : : : :
Qy      1069 GCTAACGAGCTCCTATGCTTTGTAGCGAATGTTGACGAGTAAGAGGGGAGGAT 1128      |||      : : :      |||      : : : : :
Db      216 -----IleAlaThrThrAlaGlyAlaArg----- 223      |||      : : :      |||      : : : : :
Qy      1129 GCTGCTGTTCCAGATGGGACGAGGAGTGTCTCATCTTCAACAGNAGATCCAGTA 1188      |||      : : :      |||      : : : : :
Db      223 ----- 223      |||      : : :      |||      : : : : :
Qy      1189 GTAAGTTTTTCAGAAATACTGCG---GTAGAGTTTGATGGGAACGTAGCCCGAGTAGGA 1245      : : :      |||      : : :      |||      : : : : :
Db      224 -----IleAlaAsnAsnThrGlyTyrValArgPheLeuSerAsnIleAlaSerThrSer 241      : : :      |||      : : :      |||      : : : : :
Qy      1246 GGAGGATTTACTCTCTACGGGAACGTTGCTTCTGTAATATGGAATAACCTTTGTTCTC 1305      : : :      |||      : : :      |||      : : : : :
Db      242 GlyGlyAlaIleAspAspGluGlyThrSerIleLeuSerAsnAsnLysPheLeuThrPhe 261      : : :      |||      : : :      |||      : : : : :
Qy      1306 AACATGTTGCTCTCTCTGTTTACATGCTGCTAAGCAACCAAGTGACAGGCTTCT 1365      |||      : : :      |||      : : : : :
Db      262 -----GluGlyAsnAlaAla 266      |||      : : :      |||      : : : : :
Qy      1366 AATACGAGTAATAATTACGGAGATGAGGAGCTATCTTCTGTAAGAATGTCGCGCAAGCA 1425      |||      : : :      |||      : : : : :
Db      267 LysThrThr-----GlyGlyAlaIle---CysAsnThrLysAlaSerGly 280      |||      : : :      |||      : : : : :
Qy      1426 GGA-----TCCAATACTCTGGATCAGTTTCTTCTTGTATGGAGAGGGAGTA 1470      |||      : : :      |||      : : : : :
Db      281 SerProGluLeuIleIleSerAsnAsnLys-----Thr 291      |||      : : :      |||      : : : : :
Qy      1471 GTTTCCTTTAGTACAAATGATGCTGCTGGGAAGGGGAGCTATTTATGCAAAAGCTC 1530      : : :      |||      : : :      |||      : : : : :
Db      292 LeuIlePheAlaSerAsnValAlaGluThrSerGlyGlyAlaIleHisAlaLysLysLeu 311      : : :      |||      : : :      |||      : : : : :
Qy      1531 TCGGTTGCTAACTGTGCGCTGTACAAATTTTAAAGGAATATCGCTAATGAT----- 1581      : : :      |||      : : :      |||      : : : : :
```

```
QY 2578 CGCTCTTATGTCGAGGATTAATGGTTCTCGAGTTTCGAATTTCTTCTATCATGACCGC 2637
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 649 MetGluHisGlyGlnPheTrpValSerSerMetThrAsnPheLeuHisLeuThrGly 668
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2638 GATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTATCTCTAGGAGCAACTCC 2697
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 669 AspGluAsnArgLeuGlyGlnPheArgHisThrSerGlyGlyTrpValIleGlySerAla 688
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2698 TACTTTT---GGATCATCATGATTTGGTCTGATCATTTACCGAATTTTGGTGTAGATCTAA 2754
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 689 HisThrProLysAspAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2755 GATTATGATGATGTCGTTCCATCATCATGCTGTCATAGATCGGTTTATCTATCTACC 2814
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 709 AspCysPheIleAlaHisAsnAsnSerArgThrTrpGlyGlyThrLeuPhePheLysHis 728
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2815 CAACAAGCTTTA-----TGTCGATCCTATTGTTGTCGAGATGCG 2853
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 729 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyArgAlaLysPheSerGluSer 748
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2854 TTTATC-----CGTGTAGCTACGGG 2874
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 749 AlaIleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2875 TTTGGGAATCAGCATATGAACCTCATATACATTTGTCGAGAGGAGGATGTTGCTGG 2934
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 769 HisSerAspAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTrp 788
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2935 GATAATACTGTCGCTGGAGAGATTCGAGCGGATACCGATTTGCTATTCTCCATCT 2994
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuProPheValLeuSerAsnPro 808
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2995 AAGCTCTATTGGAATGAGTTGCTGCTCTTTCGCAAGCTGAGTTTCTTATGCCGATCAT 3054
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 828
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3055 GAATCTTTACAGAGAGGCGATCAGCTCGGCATTCAGAGCGGACATCTCTAAAT 3114
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3115 CTATCAGTTCTCTGTTGGAGTGAAGTTTCATCGATGCTTCTAGTACATCTCTAAATAAT 3174
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyrThrTyr 868
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3175 AGCTTTATGGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGTACTGAGCAACG 3234
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 869 AspLeuSerGlyPhePheValSerAspValTyrArgAsnAsnProGlnSerThrAlaThr 888
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3235 CTCTCATCCCATCAAGACATGGACACAGATGCGCTTTCATTTAGCAAGCATGGAGTT 3294
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 889 LeuValMetSerProAspSerTrpLysIleArgGlyGlyAsnLeuSerArgGlnAlaPhe 908
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3295 GTGGTTAGAGGATCTATGATGCTTCTTAACAAGTATATAGAAGTATATGGCCATGGA 3354
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 909 LeuLeuArgGlySerAsnAsnTyrValTyrAsnSerAsnCysGluLeuPheGlyHisTyr 928
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3355 AGATATGATGATTCCTCTCCAGGCTATGTTGTTGAGTCGAGGAGTAGATCCGG 3414
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 929 AlaMetGluLeuArgGlySerSerArgAsnTyrAsnValAspValGlyThrLysLeuArg 948
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 3415 TTC 3417
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 949 Phe 949
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 11
D72077
polymorphic outer membrane protein g family - Chlamydophila pneumoniae (strain CWL029)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: D72077
R:Kalanen, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
```

A:Title: Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72077
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <ASN>
A:Cross-references: UNIPROT:O86164; UNIPARC:UPI000002FFER; GB:AE001628; GB:AE001363; NID
A:Experimental source: strain CWL029
C:Genetics:
C:Gene: pmp 11
C:Superfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 5.25e-67 Length: 928
Score: 1090.00 Matches: 311
Percent Similarity: 43.0% Conservative: 154
Best Local Similarity: 28.8% Mismatches: 394
Query Match: 13.8% Indels: 222
Gaps: 31

US-10-701-844-1 (1-4435) x D72077 (1-928)

```
QY 382 ATGCAACGCTCTTCCATAAGTTCTTCTTCAATGATCTAGCTTATCTTGC----- 435
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1 MetLysThrSerIleProTrpValLeuValSerSerValLeuAlaPheSerCysHisLeu 20
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 436 TGCTCTTTAAATGGGGGGGATATGCGAGAAATCATGTTCTCTCAAGGAATTCAT 495
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 21 GlnSerLeuAlaAsn-----GluGluLeuSerProAspAspSerPheAsn 36
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 496 GGGGAGACGTTAATCTGATCATTT-----CCTATACTGTTATA 534
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 37 GlyAsnIleAspSerGlyThrPheThrProLysThrSerAlaThrTyrThrLeuThr 56
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 535 GGAGATCCGAGTGGGACTACTGTTTCTGCGAGAGATTAACATTAATAAAATCTTGAC 594
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 57 GlyAsp-----ValPhePheTyrGluProGlyLysGlyThr----- 68
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 595 AATCTATGTCAGCTTTCGCTTTA-----AGTTGTTTGGGAACCTTATAGGAGTTTT 648
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 69 -----ProLeuSerAspSerCysPheLysGlnThrThrAspAsnLeu 82
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 649 ACTGTTTGGGAGGACACTCGTTCGACTTTCAGAACATACCGACTTCTACAAATGGG 708
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 83 ThrPheLeuGlyAsnGlyHisSerLeuThrPheGlyPheIleAspAlaGlyThrHisAla 102
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 709 GCAGCTCTAAGTAATAGCGCTGCTGATGGAGCTGTTTACTATTGAGGGTTTTAAAGAATTA 768
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 103 GlyAlaAlaAlaSerThrThrAlaAsnLysAsnLeuThrPheSerGlyPheSerLeuLeu 122
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 769 TCCTTTTCCAAATTCGAATTCATTTACTTCTATAGTAAATTTAGTCTCTGGAGATGGG 828
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 123 SerPheAspSer----- 126
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 829 AGCCGAGCTCCGACGACACATCTACACCGTCTAATGCTACTATTATTCTAAACAGAT 888
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 127 -----SerProSerThrThrValThrThrGlyGlnGlyThrLeuSerSerAlaGlyGly 144
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 889 CTTTGTGTACTCAATAATGAGAAGTCTCATCTCTATAGTAAATTTAGTCTCTGGAGATGGG 948
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 145 ValAsnLeuGluAsnIleArgLysLeuValAlaGlyAsnPheSerThrAlaAspGly 164
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 949 GGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCA 1008
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 165 GlyAlaIleLysGlyAlaSerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSer 184
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1009 GAAATATCTGCTCAAGCTGATGGGGGAGCTGTGTCAGTAGTCACCGATTTCTCTCTATG 1068
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 185 AsnAsnSerSerThrLysGlyGlyAla----- 194
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1069 GCTAACGAGGCTCCTATTGCTCTTTGTAGCGAATGTTGCGAGGATGAAGGGGAGGATTT 1128
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 195 -----IleAlaThrThrAlaGlyAlaArg----- 202
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
QY 1129 GCTGCTGTTTCAGGATGGGCGAGGAGTGTCTATCTTACTTCAACAGAGATCCAGTA 1188
Db 202 -----
QY 1189 GTAAGTTTTTCCAGAAATACTGCG--GTAGAGTTTGTATGGCAAGCTAGCCGAGTAGGA 1245
Db 203 -----IleAlaAsnAsnThrGlyTyrValArgPheLeuSerAsnIleAlaSerThrSer 220
QY 1246 GGAGGGATTACTCTACGGGAACGTTGCTTCTCGAATAATGGAACAACTGTTCTTC 1305
Db 221 GlyGlyAlaIleAspAspGluGlyThrSerIleLeuSerAsnAsnLysPheLeuTyrPhe 240
QY 1306 AACAAATGTTGCTTCTCTGTTTACATGCTGTGCTAAGCAACCAACAGTGCGACGCTTCT 1365
Db 241 -----GluGlyAsnAlaAla 245
QY 1366 AATACGAGTAATAATTACCGAGATGAGGAGCTATCTTCTGTAAAGAAATGTCGCGAAGCA 1425
Db 246 LysThrThr-----GlyGlyAlaIle-----CysAsnThrLysAlaSerGly 259
QY 1426 GGA-----TCCAAATACTCTGGATCAGTTTCTCTTTCATGAGAGGAGTA 1470
Db 260 SerProGluLeuIleIleSerAsnAsnLys-----Thr 270
QY 1471 GTTTCTTTAGTACGAATAGTCTGCTGGAAAGGGGAGCTATTTATGCCAAAAGCTC 1530
Db 271 LeuIlePheAlaSerAsnValAlaGluThrSerGlyGlyAlaIleHisAlaLysLysLeu 290
QY 1531 TCGGTTGCTAACTGTCGCCCTGTACAAATTTTAAAGAAATATCGCTAATGAT----- 1581
Db 291 AlaLeuSerSerGlyGlyPheThrGluPheLeuArgAsnAsnValSerSerAlaThrPro 310
QY 1582 ---GGTGGAGCGCAATTATTAGGAGAACTCGAGAGCTCAGATTATCTGCTGANTATGGA 1638
Db 311 LysGlyGlyAlaIleSerIleAspAlaSerGlyGluLeuSerLeuSerAlaGluThrGly 330
QY 1639 GATATTATTTCATGGGAATCTTAAAGAACAGCCAGCAAGAGATGCTGCCGATGTTAAT 1698
Db 331 AsnIleThrPheValArgAsnThrLeuThrThr-----GlySerThrAsp----- 346
QY 1699 GCGGTAACTGTGCTCCTCACAGCCATTTCCATGGATCGGAGCGGAAATAACACACATTA 1758
Db 347 -----ThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGluLeu 364
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTAAATGATCCCATCGAGATGGCAACGGAAAT 1818
Db 365 ArgAlaAlaLysAsnHisThrIlePhePheTyrAspProIle----- 378
QY 1819 AACCAAGCCAGGAGTCTTCCAAACTTCTAAATAATTACGATGGTAGGA----- 1869
Db 379 ThrSerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeu 398
QY 1870 -----TACACAGGGATATGTTTGTCT----- 1893
Db 399 AsnProTyrGlnGlyThrIleLeuPheSerGlyGluThrLeuThrAlaAspGluLeuLys 418
QY 1894 -----AATGGAACGAGTACTTCTTACCAAAATGTTACGATAGCAGCAAGGAGATT 1944
Db 419 ValAlaAsnAsnLeuLysSerSerPheThrGlnProValSerLeuSerGlyGlyLysLeu 438
QY 1945 GTTCTTCGTGAAGGCAAAATATACGTGAATTTCTCTAAGTCAGACAGGTGGAGTCTG 2004
Db 439 LeuLeuGlnLysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeu 458
QY 2005 TAT---ATGAGAGTCTGGAGTACATGGGATTTTGTAACTCCACAACCCACACACAGCCT 2061
Db 459 LeuGlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySer----- 473
QY 2062 CCTGCCGCTAATCAGTTGATCACCGTTTCCAAATTCGCAATTTGCTCTTTCTTCTTTGTTA 2121
Db 474 -----IleThrIleThrAsnLeuGlyIleAsnValAspSerLeuGly 487
```

```
2122 GCAAAACAATGAGTTCAGATTCGAATCTCTTACCAATCTCCAGCGCAAGATTCATCTCATCTGCA 2181
Db 488 LeuLysGlnProVal-----Ser 493
QY 2182 GTCATTGTTAGCACAACTGCTGGTCTGTTTACAAATTAGTGGCCCTATCTTTTTCAGGAT 2241
Db 494 LeuThrAlaLysGlyAlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAsp 513
QY 2242 TTGGATGATACAGCTTATGATAGGTATGCTTGGCTAGGTTCTTAATCAAAAATAATCAATGTC 2301
Db 514 IleGluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 532
QY 2302 CTGAAATTTACAGTTTAGGAGCTAAAGCCCGACAGCTAATGCCCATCATAGATTTGACTCTAGG 2361
Db 533 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 548
QY 2362 AATGAGATGCTT-----AAGTATGGCTATCAAGGAGCTGG 2397
Db 549 SerLeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlyGlnTrp 568
QY 2398 AAGCTTGGCTGGATCCTTAATACACAAATAATGCTCTTATPACTCTCGAAAGCATCATGG 2457
Db 569 AsnValAsnThrThrAspThrAlaThrAsn-----ThrLysGluAlaThrAlaThrTrp 587
QY 2458 ACTAAAACTGGGTATAATCTCGGCTCGAGCGAGTAGCTTCTTTGGTTCCAAATAGTTTA 2517
Db 588 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 607
QY 2518 TGGGATCCTATTTAGATATACGATCTGGCAATTCAGCAATTCAGCAAGTTCGATGGTGG 2577
Db 608 TrpGlyValPheThrAspIleArgSerLeuGlnGlnLeuValGluIleGlyAlaThrGly 627
QY 2578 CCGCTTATTGTCGAGGATTATGGGTTCTCGAGTTTCGAATTTCTTCTTATCATGACCGCG 2637
Db 628 MetGluHisLysGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 647
QY 2638 GATGCTTTAGTCAGGATATCGGTATATAGTGGGGTATTTCCTTAGGAGCAAACTCC 2697
Db 648 AspGluAsnArgLysGlyPheArgHisThrSerGlyGlyTyrValIleGlyGlySerAla 667
QY 2698 TACTTT---GGATCATCGATGTTGCTGATGATTTACGAAAGTATTTCGTAGATCTAA 2754
Db 668 HisThrProLysAspAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 687
QY 2755 GATTATGATGTCGTCCTTCCATCATCTGTCATAGGATCCGTTTATCTATCTATCTACC 2814
Db 688 AspCysPheIleAlaHisAsnAsnSerArgThrTyrGlyGlyThrLeuPhePheLysHis 707
QY 2815 CAACAAGCTTTA-----TGTGGATCTCTATTGTTTCGAGATGCG 2853
Db 708 SerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGlyArgAlaLysPheSerGluSer 727
QY 2854 TTTATC-----CGTCTAGCTACGGG 2874
Db 728 AlaIleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 747
QY 2875 TTTGGGAATCAGCATGATAAACTCATATACATTTTCAGAGGAGAGAGATGTCGTGG 2934
Db 748 HisSerAsnAsnArgMetGluThrHisTyrThrSerLeuProGluSerGluGlySerTrp 767
QY 2935 GATAATAACTGTCGTCGAGAGATTGGACGGGATTTACCGGATTCGATTTACTCTCATCT 2994
Db 768 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuProPheValLeuSerAsnPro 787
QY 2995 AAGCTCTATTGTAATCAGTTGCGTCTTTCGTCGAAGCTGAGTTTCTTATGCGGATCAT 3054
Db 788 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrValSerGln 807
QY 3055 GAATCTTTTACAGAGAGCGCATCAAGCTCGGGCATTCAGAGCGGACATCTCTTAAAT 3114
Db 808 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuAsn 827
QY 3115 CTATCATGTTCTCTTGGAGTGAAGTTTGTATCGATGTTTCTTAGTACACATCTCTAATAATAT 3174
```


Alignment Scores:

Pred. No.: 1.59e-66 Length: 930
 Score: 1083.00 Matches: 304
 Percent Similarity: 43.3% Conservative: 160
 Best Local Similarity: 28.4% Mismatches: 406
 Query Match: 13.7% Indels: 202
 DB: 2 Gaps: 29

US-10-701-844-1 (1-4435) x D86546 (1-930)

QY 382 ATGCAACGCTCTTCCATAAGTCTCTTCTTCAATGATCTTAGCTTATTTCTTGCTGCTCT 441
 DB |||||
 1 MetIyIleProLeuHleHisLeuLeuIleSerThrLeuVal---ThrProIleLeu 19
 QY 442 TTAATGGGGGGGATATGCGACGAAATC---ATGGTTCCTCAAGGAATTTACGATGGG 498
 DB |||||
 20 LeuSerIleAlaThrTyrglyAlaAspAlaSerLeuSerProThrAspSerPheAspGly 39
 QY 499 GAGAGTTAATGATATATCCCTATATCTGTATATAGGAGATCCGAGTGGGACTACTGTT 558
 DB |||||
 40 AlaGlyGlySerThrPheThrProLysSerThr---AlaAspAlaAsnGlyThrAsnTy 58
 QY 559 TTTCTGCGAGGAGTTAACATTAATAAATCTTGACAATTTCTATGCGAGCTTGCCCTTA 618
 DB |||||
 59 ValLeuSerGlyAsnValTyIleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly 77
 QY 619 AGTTGTTTGGGAACCTTATTAGGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACT 678
 DB |||||
 78 CysCysPheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTySerPheSer 97
 QY 679 TTCGAGAACATAGGACTTCTACAAATGGGCGAGCTCTAAGTAATVAGCGCTGCTATGGA 738
 DB |||||
 98 PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAspThrThrAlaAspLys 117
 QY 739 CTGTTTACTATGAGGTTTAAAGATATATCTCTTTCCAAATTCGAATTCATTACTTGC 798
 DB |||||
 118 AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe-----Ile 130
 QY 799 GTACTGCTGTCGCAACGACTAATAAGGGTAGCCAGACTCCGACGACACA----- 849
 DB |||||
 131 AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150
 QY 850 TCTACACCGTCTAATGGTACTATTATTTCTTAAACAGATCTTTTGTACTCAATATGAG 909
 DB |||||
 151 AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal----- 164
 QY 910 AAGTTCTCATCTATAGTAATTTACTCTCGAGATGGGGAGCTATAGTCTAAGAGC 969
 DB |||||
 165 -----SerAsnGluAlaAsnAsnGlyLysAlaIleThrAlaLysThr 179
 QY 970 TTAACGGTTCAAGGAATAGCAAGCTTTGTGTCTTCCAAAGAAATPACTGCTCAAGCTAT 1029
 DB |||||
 180 LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLysLeu 199
 QY 1030 GGGGAGCTGTGCAAGTAGTACCAAGTTCTCTGTATATGGCTTAACAGGCTCTATGCC 1089
 DB |||||
 200 GlyGlyAlaIleTySerSerAlaAlaAspIleSerGlyAsnThrGlyGlnLeuVal 219
 QY 1090 TTTGTAGCGAATGTTCAGGAGTAGAGGGGGAGGATGCTGCTGCTCAGATGGGCGAG 1149
 DB |||||
 220 PheMetAsnAsn---LysGlyGluThrGlyGlyAlaLeuGlyPheGlu----- 235
 QY 1150 CAGGAGGTGTCATCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAGAATACT 1209
 DB |||||
 236 -----AlaSerSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr 252
 QY 1210 GCGGTAGAGTTTGTATGGGAACGATGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269
 DB |||||
 253 AlaThrAsp----- 255
 QY 1270 GTTGCTTTCCTGAATAATGGAATAACCTTGTTCTTCAACAATGTTGCTTCTCTCTTAC 1329

DB 255 ----- 255
 QY 1330 ATTGCTGCTAAGCAACCAACAAAGTGGACAGGCTTCTAATACGAGTAATAATTACGAGAT 1389
 DB |||||
 256 ---AlaAla-----GlyLys 259
 QY 1390 GGAGGCTATCTTCTGTAAGAAATGGTGGCAGCAGGATCCCAATACTCTGGATCAGTT 1449
 DB |||||
 260 GlyGlyAlaIleTyCysGluLys-----ThrGlyGluThrProThrLeu 274
 QY 1450 TCCTTCATGATGAGGAGGAGTAGTTTCTTATAGTACCAATAGTCTGCGGAAAGGGGA 1509
 DB |||||
 275 ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGlyGly 294
 QY 1510 GCTATTTATGCAAAAAAGCTCTCGTGGTTGTAACCTGTGGCCTCTGACAAATTTTAAAGGAT 1569
 DB |||||
 295 AlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn 314
 QY 1570 -----ATCGCTAATGATGGTGGCGGATTTATTTAGGAGATCTGGAGAG 1614
 DB |||||
 315 ArgCysGlyAsnThrAlaAlaGlyLysGlyAlaIleAlaIleAlaAspSerGlySer 334
 QY 1615 CTGAGTTTATCTGCTGATATGAGATATTTTTCGATGGGAATCTTAAAGAACAGGCC 1674
 DB |||||
 335 LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn----- 349
 QY 1675 AAGAGAAATGCTCGCATTTAATGGCGTAATCTGTCTCTCAAGCAATTTGATGGGA 1734
 DB |||||
 350 -----ThrLeuThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyLeuGly 367
 QY 1735 TCGGAGGGAATAAAGCAATTAAGACTTAACAGCGGCATCAGATCTCTTAAATGAT 1794
 DB |||||
 368 SerSerAlaIleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyPheTyAsp 387
 QY 1795 CCCATCGAGATGGCAACGGAAAT-----AACACGCCAGCGCAG 1833
 DB |||||
 388 ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer 407
 QY 1834 TCTTCCAACTTCAAATAATTAACGATGGTGAAGGATACACAGGGATATTTGTTTGTCT 1893
 DB |||||
 408 AsnSerProLeu-----AspTySerGlyThrIleValPheSer 420
 QY 1894 -----ATGGAAGCAGTACTTTGTAC 1914
 DB |||||
 421 GlyGlyLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys 440
 QY 1915 CAAATGTTAGATAGACAGCAAGGATTTCTTCTGCGTGAAGGCAAAATTTATCAGTG 1974
 DB |||||
 441 GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLysGlyAsnValGluLeuAspVal 460
 QY 1975 AATTCTTAAGTCAGACAGGAGTCTGTATATGGAAGCTGGGAGTACATGGGATTTT 2034
 DB |||||
 461 AsnGlyPheThrGlnThrGluGlySerThrLeuLeu----- 472
 QY 2035 GTAACCTCCAAACCAACCAACAGCTCTCGCGCTAATCAGTTGATCACGCTTCCAAAT 2094
 DB |||||
 473 -----MetGlnProGlyThrLysLeuLysAlaAspThrGluAlaIleSerLeuThrLys 490
 QY 2095 CTGCAATTTGCTCTTCTTCTTGTAGCAAAACAAATGCAATGCAATTCCTCCTACCAAT 2154
 DB |||||
 491 LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer----- 505
 QY 2155 CCTCCAGCGCAAGATTTCTCATCTCGCATCTTGTGTAGCACAACTGCTGGT----- 2205
 DB |||||
 506 -----IleGluThrAlaGlyAlaAsnLys 513
 QY 2206 TCTGTTACAATTAGTGGGCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGG 2265
 DB |||||
 514 ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyGlu--- 532
 QY 2266 TATGATGGCTAGGTTCTTAATCAAAAATCAATGTCTGAAATTTACAGTTAGGAGTAAAG 2325
 DB |||||
 533 -----SerHisThrIleAsn-----GlnAlaPheThrGln 542

Db 151 AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal----- 164
Qy 910 AAGTCTCATTTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGC 969
Db 165 -----SerAsnGluAlaAsnAsnAsnGlyAlaIleThrAlaLysThr 179
Qy 970 TTAACGGTTCAAGGAATTAGCAAGCTTTGTCTCTTCCAAAGAAATACATGCTCAAGCTGAT 1029
Db 180 LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLysLeu 199
Qy 1030 GGGGAGCTGTGCAAGTAGTCACCAAGTTCTCTGCTATGGCTAAACGAGGCTCTATTGCC 1089
Db 200 GlyAlaIleIleThrSerSerAlaAlaIleSerGlyAsnThrGlyGlnLeuVal 219
Qy 1090 TTTGTAGCGAATGTTGCGAGAGTAGAAGGGGAGGAGTGTCTGCTGTTTCAGAGTGGCGAG 1149
Db 220 PheMetAsnAsn---LysGlyGluThrGlyGlyAlaLeuGlyPheGlu----- 235
Qy 1150 CAGGAGGTGCATCTACTCTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACT 1209
Db 236 -----AlaSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr 252
Qy 1210 GGGTAGAGTTTGTAGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269
Db 253 AlaThrAsp----- 255
Qy 1270 GTTGCTTTTCTGAATTAATGGAAAAACCTGTTTCTCAACAATGTTGTCTCTCTTTTAC 1329
Db 255 ----- 255
Qy 1330 ATTGCTGTAGCAACCAACAGTGGACAGGCTCTTAATACAGTAATTAATACGAGAT 1389
Db 256 ---AlaAla-----GlyLys 259
Qy 1390 GGAGAGCTATCTCTGTGAAGATGGTGGCAAGCAGGATCCAAATAACTCTGGATCAGTT 1449
Db 260 GlyGlyAlaIleIleThrCysGlyLys-----ThrGlyGluThrProThrLeu 274
Qy 1450 TCCTTTGATGGAGGGAGTAGTATTTTCTTTAGTACAAATGATGCTGGGAAAGGGGA 1509
Db 275 ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGlyGly 294
Qy 1510 GCTATTATGCCAAAAGCTCTGGTGTCTAACTGTGGCCCTGTACAAATTTTAAAGAT 1569
Db 295 AlaIleCysAlaHieGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn 314
Qy 1570 -----ATCGCTAATCATGTGGAGCGATTTATTTAGGAGAACTCGAGAG 1614
Db 315 ArgCysGlyAsnThrAlaAlaGlyLysGlyAlaIleAlaIleAlaAspSerGlySer 334
Qy 1615 CTCAGTTTATCTGCTGATTAAGAGATATTTTTCGATGGGAATCTTTAAAGAACAGCC 1674
Db 335 LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn----- 349
Qy 1675 AAAGAGATGTCGCCGATTAATGGCGTAACGTGTCTCAAGCCATTCGATGGGA 1734
Db 350 -----ThrLeuThrSerThrSerAlaProThrThrArgAsnAlaIleThrLeuGly 367
Qy 1735 TCGGAGGAGGAAATACAGACATTAAGAGCTAAAGCAGGCGATCATGATCTCTTTAATGAT 1794
Db 368 SerSerAlaLysIleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleThrPheThrAsp 387
Qy 1795 CCCATCGAGATGCCAAACGGAAAT-----AACGAGCCAGCGAG 1833
Db 388 ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer 407
Qy 1834 TCTTCCAACTTCTAAAATTAACGATGGTGAAGATACAGAGGATATGTTTGTCT 1893
Db 408 AsnSerProLeu-----AspTyrSerThrIleValPheSer 420
Qy 1894 -----AATGGAGAGAGTACTTTGTAC 1914
Db 421 GlyGlyLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys 440

Qy 1915 CAAAATGTTAGTAGAGCAAGAGGATTTGTTCTTCTGTAAGAGGCAAAATTTATCAGTG 1974
Db 441 GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLeuLysGlyAsnValGluLeuAspVal 460
Qy 1975 AATTCTCTAAGTCAGACAGGTTGGGAGTCTGTATATGGAAGCTGGGAGTACATGGATTTT 2034
Db 461 AsnGlyPheThrGlnThrGluGlySerThrLeuLeu----- 472
Qy 2035 GTAACTCCACAACCAACACAGCTCTCTCCCTCAATCACTGATTCAGCTTTGATCAGCTTTTCCAAT 2094
Db 473 -----MetGlnProGlyThrLysLeuLysAlaAspThrGluAlaIleSerLeuThrLys 490
Qy 2095 CTGCAATTTGCTCTTTCTTTCTTTAGCAAAACAATGCAAGTTAGCAATCTCTCCCTACCAAT 2154
Db 491 LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer----- 505
Qy 2155 CCTCAGCGCAAGATTCTCATCTCTGCACTCATTTGGTAGCAACAACCTGCTGTT----- 2205
Db 506 -----IleGluThrAlaGlyAlaAsnLys 513
Qy 2206 TCTGTTCAATTTAGTGGCCCTATCTTTTGGAGATTTGGATGATACAGCTTTATGATAGG 2265
Db 514 ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrGlu--- 532
Qy 2266 TATGATTTGGCTAGCTTCTAATCAAAAATCAATCTCTGTAATAATTACAGTTAGGAGCTAAG 2325
Db 533 -----SerHieThrIleAsn-----GlnAlaPheThrGln 542
Qy 2326 CCC-----CCAGCTAATGCCCATCATGATTTGACTCTAGGG----- 2361
Db 543 ProLeuValValPheThrAlaAlaThrAlaAlaSerAspIleTyrIleAspAlaLeuLeu 562
Qy 2362 -----AATGAGATCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGC 2406
Db 563 ThrSerProValGlnThrProGluProHieTyrGlyTyrGlnGlyHieTyrGluAlaThr 582
Qy 2407 TGG---GATCTAATACAGCAATAATGCTCTTATCTCTGAAAGCTACATGACTAAA 2463
Db 583 TrpAlaAspThrSerThrAlaLysSerGly-----ThrMetThrTyrValThr 598
Qy 2464 ACTGGGTATATCTGGGCGCTGAGGAGTAGCTTCTTTGTTCCAAATAGTTTATGGGA 2523
Db 599 ThrGlyTyrAsnProAsnProGluArgAlaSerValValProAspSerLeuIrpala 618
Qy 2524 TCCATTTTATAGATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGGCGCTCT 2583
Db 619 SerPheThrAspIleAerGlyThrLeuGlnGlnIleMetThrSerGlnAlaAsnSerIleTyr 638
Qy 2584 TATTGTCAGATTTATGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCGGATGCT 2643
Db 639 GlnGlnArgGlyLeuIrpAlaSerGlyThrAlaAsnPhePheHieLysAspLysSerGly 658
Qy 2644 TTAGTCAAGGATATCGTATATTAGTGGGTTTATCTCTTAGGAGCAAACTCTTAC--- 2700
Db 659 ThrAsnGlnAlaPheAerGlySerTyrGlyTyrIleValGlyGlySerAlaGluAsp 678
Qy 2701 TTTGGATCATGATTTGTTGCTAGCATTTACCGAAGTATTTGGTAGATTTATAAGATTAT 2760
Db 679 PheSerGluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysAspLeu 698
Qy 2761 GTAGTGTCTCTCCAAATCATCATGCTGTGATAGATCCGTTTATCTATCTACCCAAACA 2820
Db 699 PheIleValGluAsnThrSerHieAsnTyrLeuAlaSerLeuTyrLeuHieArgAla 718
Qy 2821 GCTTTATGTGATCCTATTTC-----TTCGGAGATGCGTTT----- 2856
Db 719 PheLeuGlyGlyLeuProMetProSerPheGlySerIleThrAspMetLeuLysAspIle 738
Qy 2857 -----ATCCGTCGTAGCTACCGTTTGGGAATCGACATATGAAACCTCA 2901
Db 739 ProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThrArg 758

```
QY 2902 TATACATTTTCAGAGGAGCGATGTTCTGTTGGGATAATAAATCTGCTGCTGAGAGATT 2961
Db 759 TyrThrSerTyrProGluAlaGlnGlySerTrpThrAsnAsnSerGlyAlaLeuGluLeu 778
QY 2962 GGAGCGGGATTACCGATTGATTTACTCCATCTTAAGCTCTATTTGAATGATGTCCTCT 3021
Db 779 GlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPheGlnGlyTyrPhePro 798
QY 3022 TTTCGTCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGCATCAA 3081
Db 799 PheLeuLysPheGlnAlaValTyrSerArgGlnGlnAsnPheLysGluSerGlyAlaGlu 818
QY 3082 GCTCGGCGCATTCAGAGCGGACATCTCCTAAATCTATCAGTTCTGTTGAGTGAAGTTT 3141
Db 819 AlaArgAlaPheAspGlyAspLeuValAsnCysSerIleProValGlyIleArgLeu 838
QY 3142 GATCGATGTTCTAGTACACATCTAATAATATAGCTTTATGGCGGCTTATATCTGTGAT 3201
Db 839 GluLysIleSerGluAspGluLysAsnAsnPheGluIleSerLeuAlaTyrIleGlyAsp 858
QY 3202 GCTTATCGGACCATCTCTGTACTGAGACAACGCTCTATCCCATCAAGAGACATGGACA 3261
Db 859 ValTyrArgLysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTrpThr 878
QY 3262 ACAGATGCGCTTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGCATCTATGATGCTTCT 3321
Db 879 SerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeuThr 898
QY 3322 CTAAACAGTAATATAGAGTATATGGCCATGGAAGATATAGTATCGAGATGCTTCTCGA 3381
Db 899 LeuSerProHisValGluLeuSerGlyGluAlaAlaTyrGluLeuArgGlySerAlaHis 918
QY 3382 GCGTATGTTGAGTCGAGGAAGTAGAGTCGGGTTTC 3417
Db 919 IleTyrAsnValAspCysGlyLeuArgTyrSerPhe 930

RESULT 15
D72078
polymorphic outer membrane protein g family - Chlamydomophila pneumoniae (strain CWL029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: D72078
R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Pan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72078
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-930 <ARN>
A:Cross-references: UNIPROT:Q92393; UNIPARC:UPI000004708C; GB:AE001627; GB:AE001363; NII
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp 8
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 3.01e-66 Length: 930
Score: 1079.00 Matches: 303
Percent Similarity: 43.2% Conservative: 160
Best Local Similarity: 28.3% Mismatches: 407
Query Match: 13.7% Indels: 202
DB: 2 Gaps: 29

US-10-701-844-1 (1-4435) x D72078 (1-930)
QY 382 ATGCMAAGCTCTTCCATAAGTCTTTCTTTCAATGATCTAGCTATTCTTGTGCTCT 441
Db 1 MetLysIleProLeuHisLeuLeuIleSerSerThrLeuVal---ThrProIleLeu 19
QY 442 TTAAATGGGGGGATATGACGACAAATC---ATGGTTCTCTCAAGGAATTTACCATGGG 498
Db 20 LeuSerIleAlaThrTyrGlyAlaAspAlaSerLeuSerProThrAspSerPheAspGly 39
```

```
QY 499 GAGACGTTAACTGTATCATTTCCCTATACTGTATTATAGGAGATCCCGAGTGGGACTACTGTT 558
Db 40 AlaGlyGlySerThrPheThrProLysSerThr---AlaAspAlaAsnGlyThrAsnTyr 58
QY 559 TTTTCTGAGGAGAGATTACATTAATAAATCTTGACAATCTTATTCAGCTTTCCTTTA 618
Db 59 ValLeuSerGlyAsnValTyrIleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly 77
QY 619 AGTTGTTTGGGAACTTATTAGGGAGCTTTTACTGTTTATAGGAGAGAGACACTCGTTGACT 678
Db 78 CysCysPheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTyrSerPheSer 97
QY 679 TTGAGAAACATACGGACTTCTCAATATGGGCGAGCTCTAAGTATATAGCGCTGCTGATGA 738
Db 98 PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAsnSerThrThrAlaAspLys 117
QY 739 CTGTTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTTGCAATTCATTCTTGCC 798
Db 118 AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe-----Ile 130
QY 799 GTACTCGCTGTCACACGACTAATAAGGGTAGCCAGACTCCGACGACAAACA----- 849
Db 131 AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150
QY 850 TCTACACCGCTCTAATGGTACTATTATTCTTAAACACAGATCTTTTGTACTCAATAATAG 909
Db 151 AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal----- 164
QY 910 AAGTTCTCATTTCTATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAGAGC 969
Db 165 -----SerAsnGluAlaAsnAsnAsnGlyGlyAlaIleThrThrLysThr 179
QY 970 TTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAAAATACTGCTCAAGCTGAT 1029
Db 180 LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLysLeu 199
QY 1030 GGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGATATGGCTAACAGGGTCTCTATTGTC 1089
Db 200 GlyGlyAlaIleTyrSerSerAlaAlaAsnSerIleSerGlyAsnThrGlyGlnLeuVal 219
QY 1090 TTTGTAGCGAATGTTGACAGATGAAGAGGGGGAGGATTCGTCTGTCAGGATGGGCGAG 1149
Db 220 PheMetAsnAsn---LysGlyGluThrGlyGlyAlaLeuGlyPheGlu----- 235
QY 1150 CAGGAGTGTCACTACTTCAACAGAGATCCAGTAGTAAGTCTTTTCCAGAAATACT 1209
Db 236 -----AlaSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr 252
QY 1210 GCGTAGAGTTTGATGGGAACGTAGCCCGTAGGAGGAGGATTTACTCTCTACGGGAAC 1269
Db 253 AlaThrAsp----- 255
QY 1270 GTTGCTTTCTGTAATAATGGAAAAACCTTGTCTTCTCAACAATGTTGCTTCTCCTGTTTAC 1329
Db 255 ----- 255
QY 1330 ATTCTGCTAAGCAACCAACAGTGGACAGGCTTCTTAATACGATAATAATTACGGAGAT 1389
Db 256 ---AlaAla-----GlyLys 259
QY 1390 GGAGGAGCTATCTCTGTAAGAATGGTGGCGACAGGATCCAAATACTCTGGATCAGTT 1449
Db 260 GlyGlyAlaIleTyrCysGluLys-----ThrGlyGluThrProThrLeu 274
QY 1450 TCCTTTGATGAGGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGTGGGAAGGGGA 1509
Db 275 ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGlyGly 294
QY 1510 GCTATTATGCAAAAAGCTCTCGTTGCTTAACGTGTCGCCCTGTACAAATTTTAAAGGAAT 1569
Db 295 AlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn 314
```


[illegible]

QY 3208 GGCACCAATCTCTGTACTGTAGACACACGCTCTCTATCCCATCAAGAGACATGGACACACGAT 326
|||
|||
Db 867 ArgAaenProLysCysGlnThrSerLeuIleAlaSerAspAlaSerMetAlaTyr 886
|||
|||
QY 3268 GCCTTTTCATTAGCAACACATGAGGTGTGGTTAGAGGATCTATGTATGCTTCTCTAACA 3327
|||
|||
Db 887 GlyThrAsnLeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsn 906
|||
|||
QY 3328 AGTAATATAGAAATATATGGCCATGGAAAGATATGAGTATCGAGATGCTTCTCGAGGCTAT 3387
|||
|||
Db 907 ProHisMetGluIlePheGlyGlnPheAlaPheGluValArgSerSerAspGlnTyr 926
|||
|||
QY 3398 GGTTCAGTCGACGAAGTAGATCCGGTTC 3417
|||
|||
Db 927 AsnThrAsnLeuGlySerLysPheCysPhe 936
|||
|||
RESULT 17
B81591
polymorphic membrane protein G family CP0308 [imported] - Chlamydia pneumoniae
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81591
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.;
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: B81591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-936 <REA>
A:Cross-references: UNIPROT:Q92898; UNIPARC:UPI0000131CED; GB:AB002193; GB:AB000
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0308
C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 1.46e-63 Length: 936
Score: 1040.00 Matches: 310
Percent Similarity: 44.1% Conservative: 162
Best Local Similarity: 29.0% Mismatches: 406
Query Match: 13.2% Indels: 192
DB: 2 Gaps: 31

US-10-701-844-1 (1-4435) x B81591 (1-936)

QY 302 ATGCAAGCTCTTCCATAAGTCTTCTTCAATGATCTAGCTTATCTTCTGCTCT 441
|||
|||
Db 1 MetLysSerSerValSerTrpLeuPhePheSerLeuPhePheSerLeuSer 20
|||
|||
QY 442 TAAATGGGGGGGATATGCAGCAGAAATCATGTT-----CCTCAAGAAATTTACGAT 495
|||
|||
Db 21 Ile-----ValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTyrAsp 36
|||
|||
QY 496 GGGGAGAGCTTAATCTGATCATCTTCCCTATCTACTGTATAGAGATCCGAGTGGGACTACT 555
|||
|||
Db 37 GlySerAsnGlyThrThrPheThrValPheSerThrThrAspAlaAlaGlyThrThr 56
|||
|||
QY 556 GTTTTTCGACGAGAGTAAACATTAATAAATCTTCACAAATCTATTCAGCGCTTGGCT 615
|||
|||
Db 57 TyrSerLeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAla 76
|||
|||
QY 616 TTAAGTTGTTTGGAACTTATTAGGAGGTTTACTCTTTTAGGAGAGGACACTCGTTG 675
|||
|||
Db 77 SerGlyCysPheLeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeu 96
|||
|||
QY 676 ACTTTCGAGACATACGGACT--TCTACAAATGGGGCAGCTCTTAAGTATATAGCGTCT 732
|||
|||
Db 97 LysPheAlaPheIleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAla 116
|||
|||
QY 733 GATGACTGTTTACTATTATGGGGTTTTAAAGAATTTATCTTTTCCAAATTCGAATTCATTA 792
|||
|||

117 AspLysAsnLeuLeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeu 136
 793 CTTGGCGTACTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGCAACATCT 852
 137 LeuLeu-----SerProThr----- 141
 853 ACACCGTCTAATGGTACTATTATCTTAAACAGACTCTTTTGTACTCAATAATGAGAAG 912
 142 -----GlyGlnCysAlaLeuLysSerValGlyAsnLeuSerLeuThrGlyAsnSerGln 159
 913 TTCTCATCTATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAA 972
 160 IleIlePheThrGlnAsnPheSerSerAspAsnGlyValIleAsnThrLysAsnPhe 179
 973 ACGTTCAAGGAATTAAGACGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGG 1032
 180 LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAla----- 195
 1033 GGNAGCTTGTCAAGTAGTACCAGTTCTCTGTATGGCTAAACGAGGCTCTATTGCTTT 1092
 195 ----- 195
 1093 GTAGCGAATGTTGCAGGAGTAAGAGGGGAGGATGCTGCTTTCAGGATGGCGACAG 1152
 196 -----PheThrGlyLysGlnGlyValValIleValIle----- 206
 1153 GGAGTGTCACTACTTCAACAGAGATCCA---GTAGTAAGTTTTCAGAAATACT 1209
 207 ---ThrGlyThrIleThrIleGluAsnSerProGlyIleValSerPheSerGlnAsnLeu 225
 1210 GCGGTAGATTGATGGGACGTAAGCCGAGTAGGAGGGAGTTTACTCTACGGGAC 1269
 226 Ala-----LysGlySer-----GlyGlyAlaLeuIleThrAspAsn 238
 1270 GTTGCTTCTGTAATATGAAACCTGTTTCTCAACAATGTGCTTCTCTGCTTTAC 1329
 239 CysSerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAla-----Trp 255
 1330 ATTGCTGCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATTACGAGAT 1389
 256 GluAlaAlaGln----- 261
 1390 GGAGGAGCTATCTTCTGTAAGAATGGTGGCCAGCAGGATCCAACTCTGGATCAGTT 1449
 262 GlyGlyAlaIleCysCys-----ThrThrAspLysThrVal 274
 1450 TCCTTTGATGAGAGGGAGTAGTTTCTTTAGTAGCAATGCTAGCTGCTGGGAAAGGGGA 1509
 275 ThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeuThrThrGlyGly 294
 1510 GCTATTTATGCCAAAAGCTCTCGGTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAAT 1569
 295 AlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeuPheGlnSerAsn 314
 1570 ATGCTAATGAT-----GGTGGAGCGATTTATTATTAGAGAAATCTGGA 1611
 315 IleSerGlySerAlaGlyGlnGlyGlyGlyAlaIleAsnIleAlaSerAlaGly 334
 1612 GAGCTCAGTTTATCTGATATTAGGAGATATTATTTTCCATGGATCTTAAAGAAACA 1671
 335 GluLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsnAsn----- 350
 1672 GCCAAGAGAATGTCGCGATGTTTAAATGGGCTAACTGTGCTCTCACAAGCCATTTTCGATG 1731
 351 -----GlnValThrAsnGlySerThrSerThrArgAsnAlaIleAsnIle 365
 1732 GGATCGGGAGGGAATAACGACATTAAAGCTTAAACAGCGGATCAGATCTCTTTAAT 1791
 366 IleAspThrAlaLysValThrSerIleArgAlaAlaThrGlyGlnSerIleThrPheThr 385
 1792 GATCCCATCGAGTGGCAACGGGAATAACGACCGGAGCTCTTCCAACTCTTAAA 1851
 |||||

386 AspProIleThr-----AsnProGlyThrAlaAlaSerThrAspThrLeuAsn 401
 1852 ATTAACGATGGTGAAGG-----TACACAGGGGATATCTGTTTTT----- 1890
 402 LeuAsnLeuAlaAspAlaAsnSerGluIleGlyThrGlyAlaIleValPheSerGly 421
 1891 -----GCTAATGGAAGCAGTAGTCTTTGTACCAA 1917
 422 GluLysLeuSerProThrGluLysAlaIleAlaAsnValThrSerThrIleArgGln 441
 1918 AATGTCATAGACAGCAAGGAGGATTTCTTCCTGAAAGCAAAATATCAGTGAT 1977
 442 ProAlaValLeuAlaArgGlyAspLeuValLeuArgPheValThrValThrPheLys 461
 1978 TCTCTAAGTCACACAGGTGGGAGT---CTGTATATGGAAGCTGGGAGTACATCGGGATTTT 2034
 462 AspLeuThrGlnSerProGlySerArgIleLeuMetAspGlyGlyThrThrLeuSer--- 480
 2035 GTAACCTCCACACACCACACAGCCTCTCTGCGCTAATTCAGTTGATCAGCTTTCCAAAT 2094
 481 -----AlaLysGluAlaAsnLeuSerLeuAsnGly 490
 2095 CTGCAATTTGCTCTTCTTCTTCTTTGTAGCAAAATGCGAGTTACGATCTCTCTACCAAT 2154
 491 LeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAla----- 504
 2155 CTTCCAGCGCAGATTTCTCATCTCTCAGTCATTGGTAGCACAACTGCTGTTCTGTATCA 2214
 505 -----AlaLeuLysThrGluAlaAlaAspLysAsnIleSer 516
 2215 ATTAGTGGCCTATCTTTTTCAGGATTTGGATGATACACACTTATGATAGTAGTTGG 2274
 517 LeuSerGlyThrIleAlaLeuIleAspThrGluGlySerPheThrGluAsnHisAsn--- 535
 2275 CTAGGTTCTAATCAAAAATCAATGTCCTGAAATTCAGTTAGGAGTACAGCCCCAGCT 2334
 536 LeuLysSerAlaSerThrTyrProLeuLeuGluLeu-----ThrThrAlaGlyAla 552
 2335 AAT-----GCCCATCAGATTTGACTCTAGGAAATGAGATGCTCTAAG 2376
 553 AsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnGluProGluThrHis 572
 2377 TATGGCTATCAAGAGAGCTGGAGCTTGGTGGATCTTAATACAGCAAAATATGCTCT 2436
 573 TyrGlyThrGlnGlyAsnTrpGlnLeuSerTrpAlaAsnAlaThrSerLys----- 590
 2437 TATCTCTGAAGACTACATGAGCTAAACTGGGTATAATCTCTGGGCTCAGCAGTAGCT 2496
 591 ---IleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluArgLysSer 609
 2497 TCTTTGTTCCAAATAGTTTATGGGATCCATTTATAGATATACGATCTGGCATTCCAGCA 2556
 610 AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu 629
 2557 ATTCAAGCAAGTGTGATGGGCTCTTATTGTCGAGGATTTAGGTTCTGAGAGTTTCG 2616
 630 IleGluThrLysSerSerGlyGluProPheGluArgLeuLeuTrpLeuSerGlyIleAla 649
 2617 AATTCTTCTATCATGACCCGATGCTTTAGGTACAGGATATCGGTATATATGATGGGGT 2676
 650 AsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly 669
 2677 TATCTCTTAGGACCAACTCTACTTT---GGATCATCGATGTTTGTGTAGCATTTACC 2733
 670 TyrAlaLeuGlyIleThrAlaThrProAlaGluAspGlnLeuThrPheAlaPheCys 699
 2734 GAAGTATTTGGTAGACTTAAAGATTATGTAGTGTCTGTTCCCAATCATCATGCTTGCATA 2793
 690 GlnLeuPheAlaArgAspAsnHisIleThrGlyLysAsnHisGlyAspThrThrGly 709
 2794 GGAATCGGTTTATCTATCTACCCCAACAGCTTTA-----TGTGGATCTTATTTGTCGGA 2847
 710 AlaSerLeuTyrPheHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly 729
 |||||

QY	2848	GATGGGTTT-----	2856
Db	730	LysAlaThrArgAlaProTrpValLeuSerGluLeSerGlnIleProLeuSerPhe	749
QY	2857	---ATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCA	2913
Db	750	AspAlaLysPheSerTyLeuHisThrAspAsnHisMetLysThrTyThr-----	767
QY	2914	GAGGAGCGGATGTCGT-----TGGGATATAACTGCTCGCTGGGAGAGATTGGAGCG	2967
Db	768	AspAsnSerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAla	787
QY	2968	GGATTACCGATTGCGATTACTCCATCTAAGCTCTATTGTAATGATGTCGCTTCCTCGTG	3027
Db	788	SerLeuProPheValIleSerValProTyLeu---LeuLysGluValGluProPheVal	806
QY	3028	CAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGCGGATCAAGCTCGG	3087
Db	807	LysValGlnTyIleTyAlaHisGlnAspPheTyGluArgTyAlaGluGlyArg	826
QY	3088	GCATTCAAGAGCGGCATCTCTAAATCTATCAGTCTCTGTTGGAGTGAAGTTTGATCGA	3147
Db	827	AlaPheAsnLysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArg	846
QY	3148	TGTTCTAGTACATCTCAATAATATATAGCTTTATGGCGGCTTATATCTGTGATCTTAT	3207
Db	847	AspSerLysSerGluLysGlyThrTyAspLeuThrLeuMetTyIleLeuAspAlaTy	866
QY	3208	CGCACCATCTCTGGTACTGAGACAGCTCCTATCCCATCAAGACATGAGCAACAGAT	3267
Db	867	ArgArgAsnProLysCysGlnThrSerLeuIleAlaSerAspAlaAsnTrpMetAlaTy	886
QY	3268	GCCTTTTCATTAGCAAGCATGGAGTCTGTTGTAGGAGTCTATGTATGCTTCTCTAACA	3327
Db	887	GlyThrAsnLeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsn	906
QY	3328	AGTATATAGAGTATATGGCCATCGAAGATATAGTATCGAGATGCTTCGAGGCTAT	3387
Db	907	ProHisMetGluIlePheGlyGlnPheAlaGluValArgSerSerArgAsnTy	926
QY	3388	GGTTTGAGTCGAGGAAGTAGTCCGGTTC 3417	
Db	927	AsnThrAsnLeuGlySerLysPheCysPhe 936	
RESULT 18			
C6546			
polymorphic outer membrane protein G family [imported] - Chlamydomphila pneumoniae (strain			
C/Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae			
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004			
C/Accession: C6546			
R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie			
Nucleic Acids Res. 28, 2311-2314, 2000			
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.			
A/Reference number: A86491; PMID:20330349; PMID:10871362			
A/Accession: C6546			
A/Status: preliminary			
A/Molecule type: DNA			
A/Residues: 1-936 <STO>			
A/Cross-references: UNIPROT:Q9Z898; UNIPARC:UPI0000131CED; GB:BA000008; NID:98978817; PI			
A/Experimental source: strain J138			
C/Genetics:			
C/Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G			
Alignment Scores:			
Pred. No.: 1.46e-63 Length: 936			
Score: 1040.00 Matches: 310			
Percent Similarity: 44.1% Conservative: 162			
Best Local Similarity: 29.0% Mismatches: 406			
Query Match: 13.2% Indels: 192			
DB: 2 Gaps: 31			

US-10-701-844-1 (1-4435) x C86546 (1-936)			
QY	382	ATGCAAACGCTTCTCCATAAGTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTCT	441
Db	1	MetLysSerValSerTrpLeuPhePheSerIleProLeuPheSerSerLeuSer	20
QY	442	TTAAATGGGGGGATATGAGCAGAAATCATGTT-----CCTCAAGGAATTTACGAT	495
Db	21	Ile-----ValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTyAsp	36
QY	496	GGGAGAGCTTTAACTGTATCATTTCCCTATCTATTAGGAGATCCGAGTGGAGTACT	555
Db	37	GlySerAsnGlyThrPheThrValPheSerThrThrAspAlaAlaAlaGlyThrThr	56
QY	556	GTTTTTTCTGAGAGAGTTAAACATTAAATAAATCTTGACAAATCTATTGACGCTTGCT	615
Db	57	TyrSerLeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAla	76
QY	616	TTAAGTTGTTTGGGAACCTTATTAGGAGTCTTCTGTTTGGGAGGAGCACTCGTTG	675
Db	77	SerGlyCysPheLeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeu	96
QY	676	ACTTTCCGAGAACATACCGACT---TCTACAAATGGGCGAGCTCTAAGTAATPAGCGTCT	732
Db	97	LysPheAlaPheIleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAla	116
QY	733	GATGAGCTGTTTACTATTAGGGTTTAAAGAATATCTCTTTTCCAAATGCAATTCATTA	792
Db	117	AspLysAsnLeuLeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeu	136
QY	793	CTTCCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAATCT	852
Db	137	LeuLeu-----SerProThr-----	141
QY	853	ACACGCTTAATGCTACTATTATTCTTAAACACAGATCTTTTGTACTCAATAATGAGAAG	912
Db	142	-----GlyGlnCysAlaLeuLysSerValGlyAsnLeuSerLeuThrGlyAsnSerGln	159
QY	913	TTCTCATCTATATTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGACTTA	972
Db	160	IleIlePheThrGlnAsnPheSerSerAspAsnGlyGlyValIleAsnThrLysAsnPhe	179
QY	973	ACGGTTCAAGAAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGCTGATGGG	1032
Db	180	LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAla-----	195
QY	1033	GGAGCTTGCTCAAGTAGTCACCCAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCC	1092
Db	195	-----	195
QY	1093	GTACGGAATGTGAGGAGTAAGAGGGGAGGGATGCTGCTGTTAGGATGGGAGCAG	1152
Db	196	-----PheThrGlyLysGlnGlyValValTyAla-----	206
QY	1153	GGAGTGTCTATCTTCTTCAACAGAGATCCA---GTAGTAAGTTTTTCCAGAAATACT	1209
Db	207	---ThrGlyThrIleThrIleGluAsnSerProGlyIleValSerPheSerGlnAsnLeu	225
QY	1210	GCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTCGGGGAAC	1269
Db	226	Ala-----LysGlySer-----GlyGlyAlaLeuTySerThrAspAsn	238
QY	1270	GTTGCTTCTTGAATAATGGAAACCTTGTCTCAACAAATGTTGCTTCTCTGTTTAC	1329
Db	239	CysSerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAla-----Tyr	255
QY	1330	ATTGCTGCTAGCAACCAACAGTGGACAGCTCTTAATACGAGTAATAATTACGAGAT	1389
Db	256	GluAlaAlaGln-----AlaGln	261
QY	1390	GGAGGAGCTATCTTCTGTAAGATGTGTGCGCAAGCAGGATCCAAATTAATCTGGAT	1449
Db	262	GlyGlyAlaIleCysCys-----ThrThrAspLysThrVal	274

QY	1450	TCCTTTGATCGAGAGGAGTAGTGTTCCTTTTAGTAGCAATGCTAGCTCTCGGAAAGGGGCA	1509
DB	275	ThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeuThrTyrGlyGly	294
QY	1510	GCTATTATGCGCAAAAGCTCTGGTGTCTAACTGTGGCCCTGTACAAATTTTAAAGAAAT	1569
DB	295	AlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeuPheGlnSerAsn	314
QY	1570	ATCGCTAATGAT-----GGTGGAGCGATTTATTATTAGGAGAAATCTGCA	1611
DB	315	IleSerGlySerSerAlaGlyGlnGlyGlyAlaIleAsnIleAlaSerAlaGly	334
QY	1612	GAGCTTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACCA	1671
DB	335	GluLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsnAsn-----	350
QY	1672	GCCAAAGAGAAATGCTGCCGATGTTAATGGCGTAACTGTGTCTCACAAAGCCATTTCCGATG	1731
DB	351	-----GlnValThrAsnGlySerThrSerThrArgAsnAlaIleAsnIle	365
QY	1732	GGATCGGGAGGAAATAACACATTAAGAGCTAAAGCAGCGGATCAGATTCCTCTTAAT	1791
DB	366	IleAspThrAlaLysValThrSerIleArgAlaAlaThrGlyGlnSerIleTyrPheTyr	385
QY	1792	GATCCCATCGAGATGGCAACGGAAATAACACAGCCAGCGAGTCTTCCAACTCTCAA	1851
DB	386	AspProIleThr-----AsnProGlyThrAlaAlaSerThrAspThrLeuAsn	401
QY	1852	ATTTAACGATGTGAAGGA-----TACACAGGGGATATTGTTTT-----	1890
DB	402	LeuAsnLeuAlaAspAlaAsnSerGluIleGluTyrGlyGlyAlaIleValPheSerGly	421
QY	1891	-----GCTAATGGAGCAGCTACTTTGTGTACCA	1917
DB	422	GluLysLeuSerProThrGluLysAlaIleAlaAsnValThrSerThrIleArgGln	441
QY	1918	AATGTTACGATAGACGAAGGAGATGTTCTTCGTGAAAGGCAAAATTTATCAGTGAAT	1977
DB	442	ProAlaValLeuAlaArgGlyAspLeuValLeuArgAspGlyValThrValThrPheLys	461
QY	1978	TCTCTAAGTCACAGAGTGGGAGT-----CTGTATATGGAAGCTGGAGTACATGGGATTTT	2034
DB	462	AspLeuThrGlnSerProGlySerArgIleLeuMetAspGlyGlyThrThrLeuSer---	480
QY	2035	GTAATCCACAACCAACCAACAGCCTCTCCCGCTAATCAGTTTGATCAAGCTTTTCCAAT	2094
DB	481	-----AlaLysGluAlaAsnLeuSerLeuAsnGly	490
QY	2095	CTGCAATTTGCTCTTTCTTTCTTTGTAGCAACAATGCGATTACGAATCCTCCTACCAAT	2154
DB	491	LeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAla-----	504
QY	2155	CCTCCAGCGCAAGATTCTCATCTCGAGCTCAATTGGTAGCAACAATGCTGTTCTGTGTACA	2214
DB	505	-----AlaLeuLysThrGluAlaAlaAspLysAsnIleSer	516
QY	2215	ATTAGTGCGCTTATCTTTTGGAGATTGGATGATACAGCTTATGATAGTATGATATGG	2274
DB	517	LeuSerGlyThrIleAlaLeuIleAspThrGluLysSerPheTyrGluAsnIleAsn---	535
QY	2275	CTAGGTTCTAAATCAAAAATCAATCTCTGAAATTTACAGTTAGGAGCTTAAGCCCCAGCT	2334
DB	536	LeuLysSerAlaSerThrTyrProLeuLeuGluLeu-----ThrThrAlaGlyAla	552
QY	2335	AAT-----GCCCATCAGATTTTGACTCTAGGGAATGAGATCCCTAAG	2376
DB	553	AsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnGluProGluThrHis	572
QY	2377	TATGGCTATCAGGAAGCTGGAGCTGGTGGGATCTTAATACAGCAAAATATGGTCTCT	2436
DB	573	TyrGlyTyrGlnGlyAsnTrpGlnLeuSerTrpAlaAsnAlaThrSerSerLys-----	590

QY	2437	TATACTCTGAAGCTACATGGACTAAACCTGGGTATAAATCTGGCGCTCAGCGAGTAGCT	2499
DB	591	---IleGlySerIleAsnIleThrArgIleGlnGlyIleProSerProGluArgLysSer	609
QY	2497	TCCTTGTGTTCCAAATAGTTTATCGGGATCCATTTTAGATATACGATCTCGGCANTCAGCA	2556
DB	610	AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu	629
QY	2557	ATTCAACGAAGTGTGGATGGCGCTCTTATTGTCGAGGATTATGGCTTCTCGAGTTTCG	2616
DB	630	IleGluThrLysSerSerGlyIleProPheGluArgGluLeuTrpLeuSerGlyIleAla	649
QY	2617	AATTCTCTCATCATGACCGCATGCTTTAGGTTCAGGGATATCGGTATATATTAGTGGGGT	2676
DB	650	AsnPhePheTyArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly	669
QY	2677	TATTCCTTAGGAGCAACTCCTACTTT---GGATCATCGATGTTTGGTCTAGCATTTACC	2733
DB	670	TyrAlaLeuGlyIleThrAlaIleThrProAlaGluAspGlnLeuThrPheAlaPheCys	689
QY	2734	GAAGTATTGGTAGATCTAAGATTATGTAGTGTGCTTCCAAATCATCATGCTTGCATA	2793
DB	690	GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLysAsnHisGlyAspThrTyrGly	709
QY	2794	CGATCCGTTTATCTATCTACCCCAACAGCTTTA-----TGTGATCTATTATTGTCGGA	2847
DB	710	AlaSerLeuTyrPheHisIleThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly	729
QY	2848	GATCGGTTT-----	2856
DB	730	LysAlaThrArgAlaProTrpValLeuSerGluIleSerGlnIleIleProLeuSerPhe	749
QY	2857	---ATCCGTGCTAGCTACGGGTGGGAAATCAGCATATGAAACCTCATATACATTTGCA	2913
DB	750	AspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrThr-----	767
QY	2914	GAGGAGCCGATGTTTCGT-----TGGGATAATAACTGTCTGGCTGGAGAGATTGGAGCG	2967
DB	768	AspAsnSerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAla	787
QY	2968	GGATTACCGATTGGATTACTCCATCTAAGCTCTATTGAATGAGTTGGCTTCCTTTCGTG	3027
DB	788	SerLeuProPheValIleSerValProTyrLeu---LeuLysGluValGluProPheVal	806
QY	3028	CAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGCATCAAGCTCGG	3087
DB	807	LysValGlnTyrIleTyrAlaHisGlnAspPheTyrGluArgTyrAlaGluGlyArg	826
QY	3088	GAATTCAGAGGGGACATCTCTAAATCTCAGTTCTCTGTGTGAGGTGAGTTGATCGCA	3147
DB	827	AlaPheAsnLysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArg	846
QY	3148	TGTTCTAGTACACATCTTAATAATATAGCTTTATGCGGCTTATATCTGTGATGCTTAT	3207
DB	847	AspSerLysSerGluLysGlyThrTyrAspLeuThrLeuMetTyrIleLeuAspAlaTyr	866
QY	3208	CGCACCATCTCTGGTACTGAGACAACGCTCTATCCCATCAAGAGACATGGAACAACAGAT	3267
DB	867	ArgArgAsnProLysCysGlnThrSerLeuIleAlaSerAspAlaAsnTrpMetAlaTyr	886
QY	3268	GCCTTTCATTAGCAAGACATGGAGTTGTGGTTAGGAGCATCTATGTTATGCTTCTTACA	3322
DB	887	GlyThrAsnLeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsn	906
QY	3328	AGTAATATAAGATGATATGGCCATGGAAGATATAGTATCGAGATGCTTCTCGAGGCTT	3388
DB	927	AsnThrAsnLeuGlySerLysPheCysPhe	936

RESULT 19

B72077

Polymorphic membrane protein G family CP0306 [imported] - Chlamydomophila pneumoniae (stra
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 R:Accession: B72077, B81592
 R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A>Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: B72077
 A:Molecule type: DNA
 A:Residues: 1-928 <ARN>
 A:Cross-references: UNIPROT:Q92398; UNIPARC:UPI0000047087; GB:AE001628; GB:AE001363; NID
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A:Reference number: AB1500; MUID:20150255; PMID:10684935
 A:Accession: B81592
 A:Molecule type: DNA
 A:Residues: 1-928 <REA>
 A:Cross-references: UNIPARC:UPI0000047087; GB:AE002192; GB:AE002161; NID:G7189226; PIDN:
 A:Experimental source: strain AR39, HL cells
 C:Genetics:
 A:Gene: pmp_9, CP0306
 C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
 Pred. No.: 2,97e-62 Length: 928
 Score: 1021.00 Matches: 313
 Percent Similarity: 44.6% Conservative: 161
 Best Local Similarity: 29.5% Mismatches: 404
 Query Match: 13.0% Indels: 184
 DB: 2 Gaps: 39

US-10-701-844-1 (1-4435) x B72077 (1-928)

QY 382 ATGCAACGCTCTTCCATAGTCTTCTTTCAATGATCTAGCTTATCTTGTGCTCT 441
 DB 1 MetLysSerLeuHisTrpPheLeuSerSerSerLeuAlaLeuPro---LeuSer 19
 QY 442 TTAATGGGGGGGATATGCGAGCAATCATGTT-----CCTCAGGAAATTAC 492
 DB 20 LeuAsnPheSerAlaPheAlaValGluIleAsnLeuGlyProThrAsnSerPhe 39
 QY 493 GATGGGGAG---ACGTTAACTGTATCATTTCCCTACTGTATAGGAGATCCGAGTGG 549
 DB 40 SerGlyProGlyThrTyThr-----ProProAlaGlnThrThrAsnAlaaspGly 56
 QY 550 ACTACTGTTTTTTCGACGAGAGTTAAATATAAAATCTTGACAAATCTTATTCAGCT 609
 DB 57 ThrIleTyAsnLeuThrGlyAspValSerIleThrAsnAla---GlySerProThrAla 75
 QY 610 TTGCTTTAAGTTTGGGAACTATTAGGAGTTTACTGTTTATAGGAGGAGGAC 669
 DB 76 LeuThrAlaSerCysPheGlyLeuThrGlyAsnLeuSerPheGlnGlyHisGlyTy 95
 QY 670 TCGTTGACTTTCGAGAACATACGGACTCTACAAATGGGCGAGCTCTAAGTAATAGCGCT 729
 DB 96 GlnPheLeuLeuGlnAsnIleaspAlaGlyAlaAsn---CysThrPheThrAsnThrAla 114
 QY 730 GCTGATGAGCTGTTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTCA 789
 DB 115 AlaAsnLysLeuLeuSerPheSerGlyPheSerTyLeuSerLeu----- 129
 QY 790 TTACTTCGGTACTCCCTGTCACACGACTAATAAGGTAGGACACTCCGACGACACA 849
 DB 130 -----IleGlnThrThrAsn 134
 QY 850 TCTACACCGCTAAATGGTACTATTATTCTTAAACACAGACTCTTTGTACTCAATATGAG 909
 DB 135 AlaThrThrGlyThrGlyAlaIleLysSerThrGlyAlaCysSerIleGlnSerAsnTy 154

QY 910 AAGTTCTCATTCTATAGTATTAGTCTCTGGAGATGGGGAGCTATAGTCTAAGAC 969
 DB 155 SerCysTyrrPheGlyGlnAsnPheSerAsnAspAnglyGlyAlaLeuGlnGlySer 174
 QY 970 TTAACGGTTCAAGAAATTAGCAAGCTTTGTGTCTTCCAGAAATATCTGCTCAAGCTGAT 1029
 DB 175 IleSerLeu---SerLeuAsnProAsnLeuThrPheAlaLysAsnLysAlaThrGlnLys 193
 QY 1030 GGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGCTTAACGAGGCTCTATTGCC 1089
 DB 194 GlyGlyAlaLeu-----TySerThr----- 200
 QY 1090 TTTGTACCGAATGTTTCAGGAGTAAGAGGGGAGGATTTGCTGCTTTCAGGATGGCAG 1149
 DB 201 -----GlyGlyIleThr----- 204
 QY 1150 CAGGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTAAAGTTTTCAGAAATACT 1209
 DB 205 -----IleAsnAsnThrLeuAsnSer-----AlaSerPheSerGluAsnThr 218
 QY 1210 GCGGTAGAGTTTGTAGGAAACGCTAGCCGAGTAGGAGGAGGATTTTACTCTCAGGAAAC 1269
 DB 219 Ala-----AlaAsnAsnGlyGlyAlaIleTyThrGluAlaSer 231
 QY 1270 GTTGCCTTCTGTAATAATGAAACCTTG---TTTCTCAACATGTTGCTTCTCCTGTT 1326
 DB 232 ---SerPheIleSerSerAsnLysAlaIleSerPheIleAsnAsn----- 245
 QY 1327 TACATTCGCTGTAAGCAACCAAGTGGACAGCTTCTAATACGAGTAATAATTACGA 1386
 DB 246 SerValThrAlaThrSerAlaThr----- 253
 QY 1387 GATGGAGGAGCTATCTTCTGTAGAATGGTGGCAAGCAGGATCCAAATAACTCTGGATCA 1446
 DB 254 ---GlyGlyAlaIleTyCys-----SerSerThrSerAlaProLysProVal 268
 QY 1447 GTTTCCTTGTAGGAGGAGGAGTAGTTTCTTTAGTAGCAATAGTAGCTGCTGGGAAAGG 1506
 DB 269 LeuThrLeuSerAspAsnGlyGluLeuAsnPheIleGlyAsnThrAlaIleThrSerGly 288
 QY 1507 GGACTTATTTATGCCAAAAGCTCGTTCGTGTAACGTGCGCTGTACAAATTTTAAAG 1566
 DB 289 GlyAlaIleTyThrAspAsnLeuValLeuSerSerGlyProThrLeuPheLysAsn 308
 QY 1567 AATATCGCTAATGAT-----GGTGGAGCGATTATTATTAGGAGATCTCGGA 1611
 DB 309 AsnSerAlaIleAspThrAlaAlaProLeuGlyGlyAlaIleAlaIleAlaAspSerGly 328
 QY 1612 GAGCTCAGTTTATCTGCTGATATGAGATATTATTTTCGATGGGATCTTAAAGAAC 1671
 DB 329 SerLeuSerLeuSerAlaLeuGlyGlyAspIleThrPheGluGlyAsn-----Thr 345
 QY 1672 GCCAAGAGATCTGCGGATGTTAATGGGTAACTGTCTCTCAAGCCATTCGATG 1731
 DB 346 ValValLysGlyAlaSer-----SerSerGlnThrThrThrArgAsnSerIleAsnIle 363
 QY 1732 GGA---TCGGAGGAGGAAATAACGACATTAAGAGCTAAAGCAGGCGATCTCTCTTT 1788
 DB 364 GlyAsnThrAsnAlaLysIleValGlnLeuArgAlaSerGlnGlyAsnThrIleTyPhe 383
 QY 1789 AATGATCCCATCGAGATGGGCAACCGAAATAACAGCCGAGGAGTCTTCCAAACTCTA 1848
 DB 384 TyrAspProIleThr-----ThrSerIleThrAlaAlaLeuSerAspAlaLeu 399
 QY 1849 AAAATTAAACGATGCGAA-----GGATACACAGCGGATATGTTTGTGCT 1893
 DB 400 AsnLeuAsnGlyProAspLeuAlaGlyAsnProAlaTyThrGlnGlyThrIleValPheSer 419
 DB 1894 -----AATGGAAGCGATCTTTGTGAC 1914
 QY 420 GlyGluLysLeuSerGluAlaGluAlaAlaGluAlaAspAsnLeuLysSerThrIleGln 439


```
QY 1915 CAATAATGTTACGATGACGACGAGGAGGATGTTCTTCGTGAAAGGCAAAATATTACAGTG 1974
|||...|||...|||...|||...|||...
Db 440 GlnProLeuThrLeuAlaGlyGlnLeuSerLeuLysSerGlyValThrLeuValAla 459
QY 1975 AATTCTCTAAGTCAGACAGGTGGAGT---CTGTATATGAAGCTGGGATCATGGGAT 2031
|||...|||...|||...|||...|||...
Db 460 LysSerPheSerGlnSerProGlySerThrLeuLeuMetAspAlaGlyThrThrLeuGlu 479
QY 2032 TTTGTAACTCCACACACACACACACGCTCTCGCGCTAACTAATCAGTTGATCGCTTCC 2091
|||...|||...|||...|||...|||...
Db 480 ThrAlaAspGly-----lleThrIleAsn 487
QY 2092 AATCTCAGTTGCTCTTCTCTCTTTGTTAGCAACAATGCAGTTACGAATCCTCTAC 2151
|||...|||...|||...|||...|||...
Db 488 AsnLeuValLeuAsnValAspSerLeu----- 496
QY 2152 AATCCTCCAGCGCAGAGATTCATCTCTGCTGAGTCATTTGGTAGCAACAATCTG 2208
|||...|||...|||...|||...|||...
Db 497 -----LysGluThrLysLysAlaThrLeuLysAlaThrGlnAlaSerGlnThr 512
QY 2209 GTTACAATTAGTGGGCTATCTTTTGGAGGATTTGGATGATACACCTTATGATGAT 2268
|||...|||...|||...|||...|||...
Db 513 ValThrLysSerGlySerLeuValAspProSerGlyAsnValTyrGluAspVal 532
QY 2269 GATTGGCTAGTTCTAATCAAAAATCAATGCTCTGAAATTACAGTTAGGACCTAAGCCC 2328
|||...|||...|||...|||...|||...
Db 533 SerTrp-----AsnAsnProGlnVal---PheSerCysLeuThrLeuThrAlaAspAsp 549
QY 2329 CCAGCTAATGCCCA---TCAGATTTTGACT-----CTAGGGAATGAGATGCTTAAG 2376
|||...|||...|||...|||...|||...
Db 550 ProAlaAsnIleHisIleThrAspLeuAlaAlaAspProLeuGluLysAsnProIleHis 569
QY 2377 TATGGCTATCAAGAGAGCTGAAGCTGCTGGATCCTTAATACAGCAATAATGCTCT 2436
|||...|||...|||...|||...|||...
Db 570 TrpGlyTyrGlnGlyAsnTrpAlaLeuSerTrpGlnGluAspThrAlaThrLysSerLys 589
QY 2437 TATACTCTGAAGCTACATGACTAAACTGGGTATAATCTCGGCTGAGCAGTAGCT 2496
|||...|||...|||...|||...|||...
Db 590 ---AlaAlaThrLeuThrTrpThrLysThrGlyTyrAsnProAsnProGluLysArgGly 608
QY 2497 TCTTTGTTCCAAATAGTTTATGGGATCCATTTAGATATACGATCTGGCATTACGCA 2556
|||...|||...|||...|||...|||...
Db 609 ThrLeuValAlaAsnThrLeuTrpGlySerPheValAspValArgSerIleGlnLeu 628
QY 2557 ATTCAAGCAAGTGTGATGGCGCTCTTATTTGTCAGGATATGGTTCTGAGTTTCG 2616
|||...|||...|||...|||...|||...
Db 629 ValAlaThrLysValArgGlnSerGlnGluThrArgGlyIleTrpCysGluGlyIleSer 648
QY 2617 AATTTCTTCTATCATGACCGGATGCTTTAGGTCAGGATATCGTATATTAGTGGGGT 2676
|||...|||...|||...|||...|||...
Db 649 AsnPhePheHisLysAspSerThrLysIleAsnLysGlyPheArgHisIleSerAlaGly 668
QY 2677 TATTCCTTAGGCAAACTCTACTTTGGATCA---TCGATGTTTGGTCTAGCATTTACC 2733
|||...|||...|||...|||...|||...
Db 669 TyrValValGlyAlaThrThrLeuAlaSerAspAsnLeuIleThrAlaAlaPheCys 688
QY 2734 GAAGTATTGTTAGATCTAAAGATATGATGTGTGCTGTTCCAAATCATCTGTCGATA 2793
|||...|||...|||...|||...|||...
Db 689 GlnLeuPheGlyLysAspArgAspHisPheIleAsnLysAsnArgAlaSerAlaTyrAla 708
QY 2794 GGATCCGTTTATCTATCTACCAACAGCTTATGTGGA-----TCTATTG 2841
|||...|||...|||...|||...|||...
Db 709 AlaSerLeuHisLeuGlnHisLeuAlaThrLeuSerSerProSerLeuLeuArgTyrLeu 728
QY 2842 TTCCGA-----GATGGTTTATCCGTGCTGCTAGCGG 2874
|||...|||...|||...|||...|||...
Db 729 ProGlySerGluSerGluGlnProValLeuPheAspAlaGlnIle-----SerTyrIle 746
QY 2875 TTGGGATCATGCATATGAAACCTCATATATTTGTCAGAGGAGAGCATGTTCTGTGG 2934
|||...|||...|||...|||...|||...
Db 747 TyrSerLysAsnThrMetLysThrTyrTyrThrGlnAlaProLysGlyGluSerSerTrp 766
QY 2935 GATTAATAACTGCTGCTGGGAGAGATTTGGACGGGATTTACCGATTGTGATTACTCCATCT 2994
```

```
Db 767 TyrAsnAspGlyCysAlaLeuGluLeuAlaSerSerLeuProHisThrAlaLeuSerHis 786
QY 2995 AAGCTCTATTGAATGAGTTGGTCTTTTCGTCGAAGCTGAGTTTCTTATGCCCCGATCAT 3054
|||...|||...|||...|||...|||...
Db 787 GluGlyLeuPheHisAlaTyrPheProHeIleIleValGluAlaSerTyrIleHisGln 806
QY 3055 GAATCTTTTACAGAGGAAGCGCATCAA---GCTCGGGCATTCAAGAGCGGACATCTCTTA 3111
|||...|||...|||...|||...|||...
Db 807 AspSerPheLysGluArgAsnThrThrLeuValArgSerPheAspSerGlyAspLeuIle 826
QY 3112 AATCTATCAGTTCCTGTTGGAGTGAAGTTTGATTCATGTTCTAGTACACATCTTAATAA 3171
|||...|||...|||...|||...|||...
Db 827 AsnValSerValProIleGlyIleThrPheGluArgPheSerArgAsnGluAlaSer 846
QY 3172 TATAGCTTTATGGCGCTTATATCTGTATCTGATGCTTATCGCACCATCTCTGTTAGTAC 3231
|||...|||...|||...|||...|||...
Db 847 TyrGluAlaThrValIleTyrValAlaAspValTyrArgLysAsnProAspCysThrThr 866
QY 3232 ACGCTCTCTATCCCATCAAGAGACATGACAGATGCTTTTCATTTAGCAAGACATGGA 3291
|||...|||...|||...|||...|||...
Db 867 AlaLeuLeuIleAsnAsnThrSerTrpLysThrThrGlyThrAsnLeuSerArgGlnAla 886
QY 3292 GTTGTGTTAGAGGATCTATGATGCTTCTCTAACAGTAATATAGAGTATATGCCCAT 3351
|||...|||...|||...|||...|||...
Db 887 GlyIleGlyArgAlaGlyIlePheTyrAlaPheSerProAsnLeuGluValThrSerAsn 906
QY 3352 GGAAGATATGATATCGAGATGCTTCTGAGGCTATGTTTGTGTCAGAGTAGTAGTC 3411
|||...|||...|||...|||...|||...
Db 907 LeuSerMetGluIleArgGlySerArgSerTyrAsnAlaAspLeuGlyGlyLysPhe 926
QY 3412 CGGTTC 3417
|||...|||...|||...|||...|||...
Db 927 GlnPhe 928

RESULT 20
E86546
polymorphic outer membrane protein G/I family [imported] - Chlamydothila pneumoniae (st
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491, MUID:20330349; PMID:10871362
A;Accession: E86546
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-928 <STO>
A;Cross-references: UNIPROT:Q9Z398; UNIPARC:UPI0000047087; GB:BA000008; NID:g8978819; P
A;Experimental source: strain J138
C;Genetics:
A;Gene: pmp_9
C;Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 2,97e-62 Length: 928
Score: 1021.00 Matches: 313
Percent Similarity: 44.6% Conservative: 161
Best Local Similarity: 29.5% Mismatches: 404
Query Match: 13.0% Indels: 184
DB: 2 Gaps: 39

US-10-701-844-1 (1-4435) x E86546 (1-928)
QY 382 ATCAACACGCTTTTCCATAGTTCTTTTCAATGATCTAGCTTATTTCTGCTGCTCT 441
|||...|||...|||...|||...|||...
Db 1 MetLysSerSerLeuHisIleTrpPheLeuIleSerSerSerLeuAlaLeuPro---LeuSer 19
QY 442 TTAATATGGGGGGGATATGACAGAGAAATCATGTT-----CCTCAGGAAATTAC 492
|||...|||...|||...|||...|||...
Db 20 LeuAsnPheSerAlaPheAlaValAlaValGluIleAsnLeuGlyProThrAsnSerPhe 39
```


Db 629 ValAlaThrLysValArgGlnSerGlnGluThrArgGlyIleTrpCysGluGlyIleSer 648
 QY 2617 AATTCTTCTCATCATGACCGCGATGCTTATAGTCACGGATATACGATATATAGTGGGCT 2676
 Db 649 AsnPhePheHisLysAspSerThrLysIleAsnLysGlyPheArgHisIleSerAlaGly 668
 QY 2677 TATTCCTTAGGACCAACTCTACTTTGGATCA---TCGATGTTTGGTCTAGCATATACC 2733
 Db 669 TyrValValGlyAlaThrThrThrLeuAlaSerAspAsnLeuIleThrAlaAlaPheCys 688
 QY 2734 GAAGTATTTGGTAGACTAAAGATTAGTAGTGCTGCTTCCATCATCATGCTTCCATA 2793
 Db 689 GlnLeuPheGlyLysAspArgAspHisPheIleAsnLysAsnArgAlaSerAlaTyrAla 708
 QY 2794 GGNATCGTTTATCTATCTACCCCAACAGCTTTATGTGGA-----TCCTATTGG 2841
 Db 709 AlaSerLeuHisLeuGlnHisLeuAlaThrLeuSerSerProSerLeuLeuArgTyrLeu 728
 QY 2842 TTCCGA-----GATCGGTTTATCCGTGCTAGTACGGG 2874
 Db 729 ProGlySerGluSerGluGlnProValLeuPheAspAlaGlnIle-----SerTyrIle 746
 QY 2875 TTTGGGAATCAGCATATGAACCTCATATACATTTTCAGAGAGAGCGGATGTCGTTGG 2934
 Db 747 TyrSerLysAsnThrMetLysThrTyrThrGlnAlaProLysGlyGlyLysSerSerTrp 766
 QY 2935 GATAATAACTGCTGCTGGAGAGATGAGCGGATACCGATTGATGATTACTCCATCT 2994
 Db 767 TyrAsnAspGlyCysAlaLeuGluLeuAlaSerSerLeuProHisThrAlaLeuSerHis 786
 QY 2995 AAGCTCTATTGATGATGTTGCTGCTTCGTCGACAGCTGATTTCTTATGCGCATCAT 3054
 Db 787 GluGlyLeuPheHisAlaTyrPheProPheIleLysValGluAlaSerTyrIleHisGln 806
 QY 3055 GAATCTTTTACAGAGAGAGCGGATCAA---GCTCGGCGCATTCAGAGCGGACATCTCTA 3111
 Db 807 AspSerPheLysGluArgAsnThrThrLeuValArgSerPheAspSerGlyAspLeuIle 826
 QY 3112 AATCTATCAGTTCTGTTGGAGTAGAGTTGATGATGCTTCTAGTACACATCTTAATAA 3171
 Db 827 AsnValSerValProIleGlyIleThrPheGluArgPheSerArgAsnGluArgAlaSer 846
 QY 3172 TATAGCTTTATGCGGCTTATCTGATGCTTATCGCACCCTCTCTGCTACTGAGACA 3231
 Db 847 TyrGluAlaThrValIleTyrValAlaAspValTyrArgLysAsnProAspCysThrThr 866
 QY 3232 ACGCTCTATCCCATCAAGAGACATGACACAGATGCCCTTTCATTTAGCAAGACATGGA 3291
 Db 867 AlaLeuLeuIleAsnAsnThrSerTyrLysThrThrGlyThrAsnLeuSerArgGlnAla 886
 QY 3292 GTTGCTGTAGAGATCTATGATGCTTCTTCAAGTATATAGATATATGCGCAT 3351
 Db 887 GlyIleGlyArgAlaGlyIlePheTyrAlaPheSerProAsnLeuGluValThrSerAsn 906
 QY 3352 GGAAGATATGATATCCAGATGCTCTCCGAGGCTATGCTTTGATGTCAGGAGTAGATC 3411
 Db 907 LeuSerMetGluIleArgGlySerSerArgSerTyrAsnAlaAspLeuGlyLysPhe 926
 QY 3412 CGGTTTC 3417
 Db 927 GlnPhe 928
 RESULT 21
 B86546
 C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: B86546
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ito
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: B86546

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1276 <STO>
 A:Cross-references: UNIPROT:Q9Z899; UNIPARC:UPI0000131CEC; GB:BA000008; NID:98978816; P:
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: pmp_6
 Alignment Scores:
 Pred. No.: 8,25e-62 Length: 1276
 Score: 1015.00 Matches: 342
 Percent Similarity: 37.9% Conservatives: 152
 Best Local Similarity: 26.3% Mismatches: 421
 Query Match: 12.9% Indels: 388
 DB: 2 Gaps: 38
 US-10-701-844-1 (1-4435) x B86546 (1-1276)
 QY 493 GATGGGAGACGTTAACTGTTATCTTATCTTATAGGAGATCCGATGGGACT 552
 Db 36 AsnGlySerSerGlySerAlaAlaPheThrAlaLysGluThrSerAspAlaSerGlyThr 55
 QY 553 ACTGTTTCTGCGAGGAGATTAACATTTAAAATCTTGACAAATCTTATTCGAGCTTTC 612
 Db 56 ThrTyrThrLeuThrSerAspValSerIleThrAsnVal---SerAlaIleThrProAla 74
 QY 613 CCTTTAAGTTGTTTGGGAACTTATTAGGAGATTTTACTGTTTATAGGAGAGACACTCG 672
 Db 75 AspLysSerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSer 94
 QY 673 TTGACTTTTCAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAAATACGCGTGT 732
 Db 95 LeuValLeuGlnThrIleAlaLeuThrHisAspGlyAlaAlaIleAsnAsnThrAsnThr 114
 QY 733 GATGGACTTTTACTATTGAGGGTTTAAAGAAATATCTTCTTCCATTTGCAATTCATTA 792
 Db 115 ---AlaLeuSerPheSerGlyPheSerSerLeu 124
 QY 793 CTT----- 795
 Db 125 LeuIleAspSerAlaProAlaThrGlyThrSerGlyLysGlyAlaIleCysValThr 144
 QY 795 --- 795
 Db 145 AsnThrGluGlyGlyThrAlaThrPheThrAspAsnAlaSerValThrLeuGlnLysAsn 164
 QY 795 --- 795
 Db 165 ThrSerGluLysAspGlyAlaAlaValSerAlaTyrSerIleAspLeuAlaLysThrThr 184
 QY 796 ---GCCGTACTGCTGCT 810
 Db 185 ThrAlaAlaLeuLeuAspGlnAsnThrSerThrLysAsnGlyGlyAlaLeuCysSerThr 204
 QY 811 GCAACGACT-----AATAAGGTTAGCAGACT---CCGACGACCAATCTTACA 855
 Db 205 AlaAsnThrValGlnGlnSerGlyThrValThrPheSerSerAsnThrAlaThr 224
 QY 856 CCGTCTAATGGTACTATTATCT----- 879
 Db 225 AspLysGlyGlyIleTyrSerLysGluLysAspSerThrLeuAspAlaAsnThrGly 244
 QY 879 --- 879
 Db 245 ValValThrPheLysSerAsnThrAlaLysThrGlyAlaTrpSerSerAspAsn 264
 QY 879 --- 879
 Db 265 LeuAlaLeuThrGlyAsnThrGlnValLeuPheGlnGluAsnLysThrThrGlySerAla 284
 QY 879 --- 879
 Db 285 AlaGlnAlaAsnAsnProGluGlyCysGlyAlaIleCysCysTyrLeuAlaThrAla 304

Qy	880	-----AAAAACAGACTCTTTTGTACTCAATAATGAGAAGTTCTCATTTCTATAGTAATTAA	933
Dd	305	ThraspyLysThrGlyLeuAlaIleSerGlnAsnGlnGluMetSerPheThrSerAsnThr	324
Qy	934	GTCTCTGGAGATGGCGAGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTCACAAG	993
Dd	325	ThrAlaAsnGlyGlyAlaIleTyAlaThrLysCysThrLeuAspGlyAsnThrThr	344
Qy	994	CTTTGTCTCTCCAAAGAAAATACTGCTCAAGCTGAT---GGCGAGCTTGCTCAAGTAGTC	1050
Dd	345	Leu---ThrPheAspGlnAsnThrAlaThrAlaGlyCysGlyAlaIleTyThrGlu	363
Qy	1051	ACC---AGTTTCTCTGATGGCTAACGAGGCTCTATTGCCTTTGPAGGAATGTGCA	1107
Dd	364	ThrGluAspPheSerLeuLysGlySerThrGlyThrValThrPheSerThrAsnThrAla	383
Qy	1108	GGAGTAGAGGGGAGGGAGT-----	1128
Dd	384	LysThr---GlyGlyAlaLeuTySerLysGlyAsnSerSerLeuThrGlyAsnThrAsn	402
Qy	1129	-----GCTGCTGTTCCAGGATGGG	1146
Dd	403	LeuLeuPheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAsnGlnGluGly	422
Qy	1147	CAGCAGGGAGTG-----TCATCATCTACTTCAACA	1176
Dd	423	CysGlyGlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysLysGlyLeu	442
Qy	1177	-----GNAGATCCAGTAGTAGTATTTCCAGAAATCTGG	1212
Dd	443	TripLeuAspAsnGluAsnValSerLeuSerGlyAsnThrAlaThrValSerGlyGly	462
Qy	1213	-----GTAGAGTTTTGATGGCAAC	1230
Dd	463	AlaIleTyAlaThrLysCysAlaLeuHisGlyAsnThrThrLeuThrPheAspGlyAsn	482
Qy	1231	GTAGCCGA---GTAGAGGAGGGATTTAC-----	1257
Dd	483	ThrAlaGluThrAlaGlyGlyAlaIleTyThrGluThrGluAspPheThrLeuThrGly	502
Qy	1258	TCCTACCGGAACGTTGCTTCTCCTCGAATAAT-----	1287
Dd	503	SerThrGlyThrValThrPheSerThrAsnThrAlalysThrAlaGlyAlaLeuHisThr	522
Qy	1288	---GGAAAAACCTTGTTTCTCAACAATGTGCTTCTCCTGTTTACATTGCTGCTAAGCAA	1344
Dd	523	LysGlyAsnThrSerPheThrLysAsnLysAla-----LeuValPheSerGlyAsn	539
Qy	1345	CCAACAATGGACAGGCTTCTAATACGATGATTAATTAACGAGATGAGGAGCTATCTTC	1404
Dd	540	SerAlaThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyAlaIleLeu	559
Qy	1405	TGT-----AGAAATGGTGGC	1419
Dd	560	CysAsnIleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAsnGluSer	579
Qy	1420	CAAGCAGGATCCAAATACTCTGGATCAGTTTCC-----	1452
Dd	580	LeuSerPheIleAsnAsnThrAlaLysArgSerGlyGlyIleTyAlaProLysCys	599
Qy	1453	---TTTGATGGAGGGNGTAGTTTTCTTTTAGTAGCAATGATGCTGCTGGGAAGGGGA	1509
Dd	600	ValIleSerGlySerGluSerIleAsnPheAspGlyAsnThrAlaGluThrSerGlyGly	619
Qy	1510	GCTATTTATGCAAAAAGCTCTCGGTTGCTTAACGTGTGCCCTGTACAAATTTTAAAGNAT	1569
Dd	620	AlaIleTySerLysAsnLeuSerIleThrAlaAsnGlyProValSerPheThrAsnAsn	639
Qy	1570	ATCGTAAATGATGGAGCAATTTATTAGAGAAATCTGAGAGCTCAGTTTATCTGCT	1629
Dd	640	SerGlyGlyLysGlyAlaIleTyAlaIleAlaSerGlyGlyLeuSerLeuGluAla	659

1630	Qy	GATTATGGAGATATATTTTCGATGGGAATCTTTAAAGAACACAGCCAAAGAGATGCTGCC	1689
660	Db	IleAspGlyAspIleThrPheSerGlyAsn-----ArgAlaThrGluGlyThrSer	676
1690	Qy	GATGTTAATGGCTAACTGTGTCTCACAAGCCATTTTCGATGGATCGGAGGGAATAA	1749
677	Db	ThrProAsn-----SerIleHisLeuGlyAlaGlyAlaLysIle	689
1750	Qy	ACGACATTAAAGAGCTAAAGCAGGGCATCAGATTTCTTTTAATCATCCCATCGAGATG---	1806
690	Db	ThrLysLeuAlaAlaProGlyHisThrIleTyrPheTyrAspProIleThrMetGlu	709
1807	Qy	-----GCAAAACCGA-----	1815
710	Db	AlaProAlaSerGlyGlyThrIleGluGluLeuValIleAsnProValVallysAlaIle	729
1816	Qy	-----ANTAACACGACGAGCGACGCTCTCCAAACTTCTTAAAAATTAAC	1857
730	Db	ValProProGlnProLysAsnGlyProIleAlaSerValProValValProValAla	749
1858	Qy	GATGGTGAAGCATACACAGGGGATATTGTTTTT-----	1890
750	Db	ProAlaAsnProAsnThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAsp	769
1891	Qy	-----GCTAATGGAAGCAGCTACTTTGTACAAAATGTTACGATAGACGAAGGA	1938
770	Db	AlaSerIleProAlaAsnThrThrIleLeuAsnGlnLysIleAsnLeuAlaGlyGly	789
1939	Qy	AGGATTTGTTCTCGTGAAGGCAAAATATATCATGTGAATTCCTCTAAGTCAGACAGGTGGG	1998
790	Db	AsnValValLeuLysGluGlyAlaThrLeuGlnValTyrSerPheThrGlnGlnProAsp	809
1999	Qy	AGT---CTGTATATGGAAGCTGGAGTACATGGATTTTGTAACTCCACAACCAACCA	2055
810	Db	SerThrValPheMetAspAlaGlyThrThrLeuGluThrThrThrThrAsn-----	826
2056	Qy	CAGCTCTCCCGCTAATCATAGTTGATCAGCTTTTCCAAATCGCATTTGTCCTCTTTCT	2115
827	Db	-----AsnThrAspGlySerIleAspLeuLysAsnLeuSerValAsnLeuAspAla	843
2116	Qy	TTGTTAGCAAAACATGCAGTTACGAATCTCTCCAAATCCTCCAGCGCAAGATTCTCAT	2175
844	Db	LeuAspGlyLysArgMetIleThr-----	851
2176	Qy	CCTGCAGTCATTGTTAGCACACTGCTGGTTCTGTTACAAATTAGTGGCGCTATCTTTT	2235
852	Db	-----IleAlaValAsnSerThrSerGlyLysLysIleSerGlyAspLeuLysPhe	869
2236	Qy	GAGGATTTGGATGATACAGCTTATGATAGGTATGATTTGGCTAGGTTCTTAATCAAAAATC	2295
870	Db	HisAsnAsnGluGlySerPheTyrAspAsn-----ProGlyLeuLysAlaAsnLeu	886
2296	Qy	AATGCTCTGAAATTCAGATTA---GGGACTAAGCCCCCAGCTAAT-----	2337
887	Db	AsnLeuProPheLeuAspLeuSerSerThrSerGlyThrValAsnLeuAspAspPheAsn	906
2338	Qy	---GCCCCATCAGATTTGACTCTAGGGAATGAGATGCTAAGTATGCTATCAAGGAAGC	2394
907	Db	ProIleProSerSerMetAla-----AlaProAspTyrGlyTyrGlnGlySer	922
2395	Qy	TGGAAGCTTGGTGGGATCTTAATACAGCAAAATAATGTCCTTATATCTCTGAAAGCTACA	2454
923	Db	TrpThrLeuVal-----ProLysValGlyAlaGlyLysValThrLeuValAlaGlu	940
2455	Qy	TGGACTAAACCTGGGTATATCTTGGGCGCTGAGGAGTAGCTTTTGGTTCCAAATAGT	2514
941	Db	TrpGlnAlaLeuGlyTyrThrProLysProGluLeuArgAlaThrLeuValProAsnSer	960
2515	Qy	TTATGGGATCCATTTTAGATATACGATCTGGCATTCAGCAATTCAGCAAGTGTGAT	2574
961	Db	LeuTrpAsnAlaTyrValAsnIleHisSerIleGlnGlnGluIleAlaThrAlaMetSer	980
2575	Qy	GGGCGCTCTTATGTCGAGGATTTATGGGTTTCTGGAGTTTCGAATTTCTTCTATCATGAC	2634

981 AspAlaProSerHisProGlyIleThrPheGlyGlyLeuAlaPheHisGlnAsp 1000
 2635 CCGCATGCTTTAGTCAGGATATCGGTATATAGTGGGGTTATTCCTAGGAGCAAC 2694
 1001 LysGlnLysGluAsnAlaGlyPheArgLeuIleSerArgGlyTyrIleValGlyLys 1020
 2695 TCTACTTTGGATCATCG--ATGTTTGTCTAGCATTTACCGAAGTATTCGTAGACT 2751
 1021 MetThrThrProGlnGluTyrThrPheAlaValAlaPheSerGlnLeuPheGlyLys 1040
 2752 AAGATTATGATGTGTCGTTCCATCATCATGCTTGATAGATCCGCTTATCTATCT 2811
 1041 LysAspTyrValValSerAspIleLysSerGlnValTyrAlaGlySer-----Leu 1058
 2812 ACCCAACAAGCTTTATGTGATCCTATTTGTCGGAGATCGGTTATCCGTGCT----- 2865
 1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgHisValLeuSer 1078
 2866 -----AGCTACGG 2873
 1078 sValLeuProGluLeuProGlyGluThrProLeuValLeuHisGlyLysValSerTyr 1098
 2874 GTTTGGGAATCAGCATATGAAACCTCATATACATTTGACAGAGAGCGATGTCGTTG 2933
 1098 YargAsnHisHisAsnMetThrThrLysLeuAlaAsnAsnThrGlnGlyLysSerAsp 1118
 2934 GGATAATACTGTCGCTGAGAGATGAGCGGATACCGATTGATGATCTCATC 2993
 1118 PaspSerHisSerPheAlaValGluValGlyGlySerLeuProValAspLeu-----As 1136
 2994 TAAGCTCTATTGAATGAGTGGCTGCTTCGTCGAGCTGAGTTTCTTATCCCATCA 3053
 1136 nTyrArgTyrLeuThrSerTyrSerProTyrValLysLeuGlnValValSerVal 1156
 3054 TGAATCTTTACAGAGAGAGCGCATCAAGCTCGGCATTCAGAGCGGACATCTCTAAA 3113
 1156 nLysGlyPheGlnGluValAlaAlaAspProArgIlePheAspAlaSerHisLeuVal 1176
 3114 TCTATCAGTTCCTGTTGGAGTGAAGTTTGATCATGTTCTAGTACATCTCTAATAA 3173
 1176 nValSerIleProMetGlyLeuThrPheLysHisGluSerAlaLysProSerAlaLe 1196
 3174 TAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTTACTGAGACA 3233
 1196 uLeuLeuThrLeuGlyTyrAlaValAspAlaTyrArgAspHisProHisCysLeuThr 1216
 3234 GCTCCTATCCATCAAGAGACATGGACAACAGATGCTCTTCATTTAGCAAGACATGG 3293
 1216 r---LeuThrAsnGlyThrSerTyrPheAlaThrAsnLeuSerArgGlnAlaPhe 1235
 3294 TGTGTTAGAGGATCTATGATGCTCTCTAAACAAGTAATATAGAGTATATGCGCATGG 3353
 1235 ePheAlaGluAlaSerGlyHisLeuLysLeuHisGlyLeuAspCysPheAlaSerG 1255
 3354 AAGATATGATATCGAGATGCTTCTCGAGGCTATGTTGATGTCGAGAGATGAGTCCG 3413
 1255 ySerCysGluLeuArgSerSerArgSerTyrAsnAlaAsnCysGlyThrArgTyrSe 1275
 3414 GTTC 3417
 1275 rPhe 1276

RESULT 22

C81591
 Polymorphic membrane protein G family CP0309 [imported] - Chlamydomonada pneumoniae (stra
 C/Species: Chlamydomonada pneumoniae, Chlamydia pneumoniae
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C/Accession: C81591
 R/Read: T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 , C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: C81591
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1276 <REA>
 A:Cross-references: UNIPROT:Q92899; UNIPARC:UPI0000131CEC; GB:AE002161; N1
 A:Experimental source: strain AR39, HL cells
 C:Genetics:
 A:Gene: CP0309
 Alignment Scores:
 Pred. No.: 8,25e-62 Length: 1276
 Score: 1015.00 Matches: 342
 Percent Similarity: 37.9% Conservative: 152
 Best Local Similarity: 26.3% Mismatches: 421
 Query Match: 12.9% Indels: 388
 DB: 2 Gaps: 38

US-10-701-844-1 (1-4435) x C81591 (1-1276)

QY 493 GATGGGAGAGCTTAACGTATCTTCCCTATCTGTTATAGGAGATCCGAGTGGACT 552
 Db 36 AsnGlySerSerGlySerAlaAlaPheThrAlaLysGluThrSerAspAlaSerGlyThr 55
 QY 553 ACTGTTTTCTCGAGGAGATTAAACATTAAAAAATCTTGACAATTTATTGCGAGTTTG 612
 Db 56 ThrTyrThrLeuThrSerAspValSerIleThrAsnVal---SerAlaIleThrProAla 74
 QY 613 CTTTAAAGTTGTTGGGAACCTTATTAGGAGTTTACTGTTTACGGAGAGAGACACTCG 672
 Db 75 AspLysSerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSer 94
 QY 673 TTGACTTTCAGAACATACGACTCTACAAATGGGCGAGCTCTAAGTAAATAGCGCTGCT 732
 Db 95 LeuValLeuGlnThrIleAlaLeuThrHisAspGlyAlaAlaIleAsnAsnThrAsnThr 114
 QY 733 GATGGACTGTTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTTGCAATTCATTA 792
 Db 115 -----AlaLeuSerPheSerGlyPheSerSerLeu 124
 QY 793 CTT----- 795
 Db 125 LeuIleAspSerAlaProAlaThrGlyThrSerGlyGlyLysGlyAlaIleCysValThr 144
 QY 795 ----- 795
 Db 145 AsnThrGluGlyGlyThrAlaThrPheThrAspAsnAlaSerValThrLeuGlnLysAsn 164
 QY 795 ----- 795
 Db 165 ThrSerGluLysAspGlyAlaAlaValSerAlaTyrSerIleAspLeuAlaLysThrThr 184
 QY 796 -----CCGACTGCTGCTGCT 810
 Db 185 ThrAlaAlaLeuLeuAspGlnAsnThrSerThrLysAsnGlyGlyAlaLeuCysSerThr 204
 QY 811 GCACAGACT-----ATAAGGTAGCCAGACT---CCGACGACACATCTACA 855
 Db 205 AlaAsnThrThrValGlnGlyAsnSerGlyThrValThrPheSerSerAsnThrAlaThr 224
 QY 856 CCGTCTAATGGTACTATTATTCT----- 879
 Db 225 AspLysGlyGlyGlyIleThrLysGluLysAspSerThrLeuAspAlaAsnThrGly 244
 QY 879 ----- 879
 Db 245 ValValThrPheLysSerAsnThrAlaLysThrGlyGlyAlaTyrSerSerAspAsn 264
 QY 879 ----- 879
 Db 265 LeuAlaLeuThrGlyAsnThrGlnValLeuPheGlnGluAsnLysThrThrGlySerAla 284
 QY 879 ----- 879

Db 285 AlaGlnAlaAsnAsnProGluGlyCysGlyGlyAlaIleCysCysThrLeuAlaThrAla 304
QY 880 -----AAAACAGATCTTTGTTACTCAATATGAGAAAGTTCTCATTTCTATAGTAATTTA 933
Db 305 ThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMetSerPheThrSerAsnThr 324
QY 934 GTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAGAGGAATTAAGCAAG 993
Db 325 ThrThrAlaAsnGlyGlyAlaIleThrAlaThrLysCysThrLeuAspGlyAsnThrThr 344
QY 994 CTTTGTGCTCCCAAGAAATACTGCTCAAGCTGAT---GGGGAGCTTGTCAGAGTAGTC 1050
Db 345 Leu---ThrPheAspGlnAsnThrAlaThrAlaGlyCysGlyGlyAlaIleThrThrGlu 363
QY 1051 ACC---AGTTTCTCTGCTATGCTTAACAGGCTCTATTGCTTTGTAGCGAATTTGCA 1107
Db 364 ThrGluAspPheSerLeuLysGlySerThrGlyThrValThrPheSerThrAsnThrAla 383
QY 1108 GGAGTAAGAGGGGGAGGAT--- 1128
Db 384 LysThr---GlyGlyAlaLeuThrSerLysGlyAsnSerSerLeuThrGlyAsnThrAsn 402
QY 1129 -----GCTGCTGTTCCAGATCGG 1146
Db 403 LeuLeuPheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAsnGlnGluGly 422
QY 1147 CAGCAGGGAGTG---TCATCATCTACTTCAACA- 1176
Db 423 CysGlyGlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysLysGlyLeu 442
QY 1177 -----GAAGATCCAGTAGTAAGTTTCCAGAAATACTGCG- 1212
Db 443 TrpIleGluAspAsnGluAsnValSerSerGlyAsnThrAlaThrValSerGlyGly 462
QY 1213 -----GTAGAGTTTCATGGGAAC 1230
Db 463 AlaIleThrAlaThrLysCysAlaLeuHisGlyAsnThrThrLeuThrPheAspGlyAsn 482
QY 1231 GTAGCCCGA---GTAGGAGGGAGGATTTAC- 1257
Db 483 ThrAlaGluThrAlaGlyGlyAlaIleThrGluThrGluAspPheThrLeuThrGly 502
QY 1258 TCCTACGGGAACGTTGCTTCTGTAATAAT- 1287
Db 503 SerThrGlyThrValThrPheSerThrAsnThrAlaLysThrAlaGlyAlaLeuHisThr 522
QY 1288 ---GGAAAAACCTTTCTCAACATGTTGCTTCTCTGTTTACATGCTGCTAAGCAA 1344
Db 523 LysGlyAsnThrSerPheThrLysAsnLysAla-----LeuValPheSerGlyAsn 539
QY 1345 CCAACAAGTGGACAGCTTCTAATACGAGTAATAATTACGGAGATGGAGGAGCTATCTTC 1404
Db 540 SerAlaThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyGlyAlaIleLeu 559
QY 1405 TGT-----AAGAATGGTCG 1419
Db 560 CysAsnIleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAsnGluSer 579
QY 1420 CNAAGCAGGATCCAATAACTCTGGATCTTCC- 1452
Db 580 LeuSerPheIleAsnAsnThrAlaLysArgSerGlyGlyIleThrAlaProLysCys 599
QY 1453 ---TTTGATGGAGAGGAGTAGTTTCTTTAGTCAATGTAGCTGCTGGGAAAGGGGA 1509
Db 600 ValIleSerGlySerGluSerIleAsnPheAspGlyAsnThrAlaGluThrSerGlyGly 619
QY 1510 GCTATTATGCCAAAAGCTCTCGTGTCTAATCTGTGGCCCTGTGACAAATTTTAAGGAAT 1569
Db 620 AlaIleThrSerLysAsnLeuSerIleThrAlaAsnGlyProValSerPheThrAsnAsn 639
QY 1570 ATCGCTAATGATGGTGGAGGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCT 1629

Db 640 SerGlyGlyLysGlyGlyAlaIleThrIleAlaAspSerGlyGluLeuSerLeuGluAla 659
QY 1630 GATTATGAGATATATTATTTTCGATGGGAATCTTAAAGAACAGCAAGAGAATGCTGCC 1689
Db 660 IleAspGlyAspIleThrPheSerGlyAsn-----ArgAlaThrGluGlyThrSer 676
QY 1690 GATGTTAATGGCTAACTGTGCTCTACAAGCCATTTTCGATGGATCGGAGGGGAAATA 1749
Db 677 ThrProAsn-----SerIleHisLeuGlyAlaGlyAlaLysIle 689
QY 1750 ACGACATTAAGAGCTAAAGCAGGCGCATCAGATTTCTTTAATGATCCCATCGAATG--- 1806
Db 690 ThrLysLeuAlaAlaProGlyHisThrIleThrPheThrAspProIleThrMetGlu 709
QY 1807 -----GCAACCGA- 1815
Db 710 AlaProAlaSerGlyGlyThrIleGluGluLeuValIleAsnProValValLysAlaIle 729
QY 1816 -----AATAACCGCCAGCGCAGTCTTCCAAACTTCTAAAAATTAAAC 1857
Db 730 ValProProGlnProLysAsnGlyProIleAlaSerValProValProValAla 749
QY 1858 GATGCTGAAGGATACACAGGGGATATTGTTTT- 1890
Db 750 ProAlaAsnProAsnThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAsp 769
QY 1891 -----GCTAATGAAGCAGTACTTTGTACCAAAATGTTAGATAGACAGGA 1938
Db 770 AlaSerIleProAlaAsnThrThrThrIleLeuAsnGlnLysIleAsnLeuAlaGlyGly 789
QY 1939 AGGATTTCTTCTGTAAGGCAAAATTTATCATGTAATCTCTAAGTCACACAGGTGG 1998
Db 790 AsnValValLeuLysGlyAlaThrLeuGlnValThrSerPheThrGlnGlnProAsp 809
QY 1999 AGT---CTGTATATGAAGCTGGGAGTACATGGGATTTGTAACTCCACACACCACAA 2055
Db 826 810 SerThrValPheMetAspAlaGlyThrThrLeuGluThrThrThrAsn----- 826
QY 2056 CAGCCTCTCGCCGTAATCAGTTGATCAGCTTTCCAAATCTGCATTTGTCTTCTTCTCT 2115
Db 827 -----AsnThrAspGlySerIleAspLeuLysAsnLeuSerValAsnLeuAspAla 843
QY 2116 TTGTAGCAACAATGAGTTCAGATTCGAATCTCTCACTCAATCTCCAGCGCAAGATCTCAT 2175
Db 844 LeuAspGlyLysArgMetIleThr- 851
QY 2176 CTTGACGATCTGTTAGCACAACTGCTGTTCTGTACAAATAGTGGGCTTATCTTTT 2235
Db 852 -----IleAlaValAsnSerThrSerGlyLysLysIleSerGlyAspLeuLysPhe 869
QY 2236 GAGGATTTGGATGATACAGCTTATGATAGGTATGATTTGGCTAGTCTTAATCAAAAATC 2295
Db 870 HisAsnAsnGluGlySerPheThrAspAsn-----ProGlyLeuLysAlaAsnLeu 886
QY 2296 AATGCTCGAAATTAACAGTTA---GGGACTAAGCCCCCAGCTAAT- 2337
Db 887 AsnLeuProPheLeuAspLeuSerSerThrSerGlyThrValAsnLeuAspPheAsn 906
QY 2338 ---GCCCATCAGATTTGACTCTAGGGAATGAGATGCTAAGTATGGCTATCAAGGAAGC 2394
Db 907 ProIleProSerSerMetAla-----AlaProAspThrGlyThrGlnGlySer 922
QY 2395 TGGAGCTTGGTGGGATCTTAATACAGCAAAATATGTCCTTATCTTCTCAAAAGCTACA 2454
Db 923 TrpThrLeuVal-----ProLysValGlyAlaGlyLysValThrLeuValAlaGlu 940
QY 2455 TGGACTAAACTGGGTATAATCTGGGCTCAGCGAGTAGTCTTCTTGGTCCAAATAGT 2514
Db 941 TrpGlnAlaLeuGlyThrThrProLysProGluLeuArgAlaThrLeuValProAsnSer 960
QY 2515 TTATGGGATCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGGTGTGAT 2574
Db 961 LeuTrpAsnAlaThrValAsnIleHisSerIleGlnGlnIleAlaThrAlaMetSer 980

QY 2575 GGGCGCTCTATTGTGCGAGGATTATGGGTTTCTGGAGTTTCGAATTTCTTCTATCATGAC 2634
Db 981 AspaIaProSerHisProGlyIleIlePheGlyGlyIleGlyAsnAlaPheHisGlnAsp 1000
QY 2635 CCGCATGCTTTAGGTGAGGATATCGGTATATAGTGGGGTTATCTTACGAGCAAC 2694
Db 1001 LysGlnLysGluAsnAlaGlyPheArgLeuIleSerArgGlyTyxIleValGlyGlySer 1020
QY 2695 TCTACTTTGGATCATCG--ATGTTTGGTCTAGCAATTTACCGAAGTATTTGGTAGACT 2751
Db 1021 MetThrThrProGlnGluTyxThrPheAlaValAlaPheSerGlnLeuPheGlyLysSer 1040
QY 2752 AAGATTTATGTAGTGTGCTTCCATCATCATGCTTGCATAGATCCGTTTATCTATCT 2811
Db 1041 LysAspTyxValValSerAspIleLysSerGlnValTyxAlaGlySer-----LeuCys 1058
QY 2812 ACCCAACAAGCTTTATGTGGATCTTATTTGTCGGAGATCGGTTTATCCGTGCT----- 2865
Db 1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgArgHisValLeuSerLys 1078
QY 2866 -----AGCTACGG 2873
Db 1078 sValLeuProGluLeuProGlyGluThrProLeuValLeuHisGlyGlnValSerTyxG 1098
QY 2874 GTTTCGGAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGATGTTGTTG 2933
Db 1098 yArgAsnHisHisAsnMetThrThrLysLeuAlaAsnAsnThrGlnGlyLysSerAspTr 1118
QY 2934 GATATAAATGCTGCTCGCGAGAGATGAGCGGATTTACCGATTGTGATTTACTCCATC 2993
Db 1118 pAspSerHisSerPheAlaValGluValGlyGlySerLeuProValAspLeu-----As 1136
QY 2994 TAAGTCTATTGAATGAGTGTGCTCTTTCGTCGCAAGCTGATTTTCTTATCCCATCA 3053
Db 1136 nTyxArgTyxLeuThrSerTyxSerProTyxValLysLeuGlnValValSerValAsnG 1156
QY 3054 TGAATCTTTACAGAGGAGCGCATCAAGCTCGGCATTTCAAGAGCGGACATCTCTTAA 3113
Db 1156 nLysGlyPheGlnGluValAlaAlaAspProArgIlePheAspAlaSerHisLeuValAs 1176
QY 3114 TCTATCAGTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCTCTTAATAA 3173
Db 1176 nValSerIleProMetGlyLeuThrPheLysHisGluSerAlaLysProProSerAlaLe 1196
QY 3174 TAGCTTTATGCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTTACTGAGCA 3233
Db 1196 uLeuLeuThrLeuGlyTyxAlaValAlaAspAlaTyxArgAspHisProHisCysLeuThrSe 1216
QY 3234 GCTCCTATCCATCAGAGATCGGACACAGATGCTTTCATTTACGAGCATGGAGT 3293
Db 1216 r---LeuThrAsnGlyThrSerTrpSerThrPheAlaThrAsnLeuSerArgGlnAlaPh 1235
QY 3294 TGTGTTAGAGGATCTATGATGCTTCTCTAAACAAGTAATATAGATATATGCGCATGG 3353
Db 1235 ePheAlaGluAlaSerGlyHisLeuLysLeuLeuHisGlyLeuAspCysPheAlaSerG 1255
QY 3354 AAGATATGAGTATCGAGATCTTCTCGAGGCTATGTTTGCAGTCGACGAAGTATGAGTCG 3413
Db 1255 ySerCysGluLeuArgSerSerSerArgSerTyxAsnAlaAsnCysGlyTyxArgTyxSe 1275
QY 3414 GTTC 3417
Db 1275 rPhe 1276

RESULT 23

B72078
Polymorphic outer membrane protein g/i family - Chlamydophila pneumoniae (strain CWL029)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: B72078
R;Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999

A;Title: Comparative Genomes of Clamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: B72078

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-1407 <ARN>

A;Cross-references: UNIPROT:Q92899; UNIPARC:UPI0000047086; GB:AE001363; NIT

A;Experimental source: strain CWL029

C;Genetics:

A;Gene: pmp_6

Alignment Scores:
Pred. No.: 3.5e-60 Length: 1407
Score: 991.50 Matches: 334
Percent Similarity: 40.3% Conservative: 150
Best Local Similarity: 27.8% Mismatches: 440
Query Match: 12.6% Indels: 277
DB: 2 Gaps: 40

US-10-701-844-1 (1-4435) x B72078 (1-1407)

QY 376 TGTGTGATGCAACGCTCTTTCCATAGTTCCTTTCTTTCA-----ATG 417
Db 299 CysTyxLeuAlaThrAlaThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMet 318
QY 418 ATTCTAGCTTATCTTGTCTCTTTAAATGGGGGGA---TATGCAGCAGAAATCATG 474
Db 319 SerPheThrSerAsnThrThrThrAlaAsnGlyGlyAlaIleTyxAlaThrLysCysThr 338
QY 475 GTTCTCTCAAGGAATTACATGGGAGAGCTTAATCTGTATCATTTCCCTATATCTGTATA 534
Db 339 Leu-----AspGlyAsnThr---ThrLeuThrPheAspGlnAsnThrAla 352
QY 535 GGAGATCCGAGTGGGACTACTCTTTTCTGCGAGGAGCTTAACATTAAAAATCTTGAC 594
Db 353 ThrAlaGlyCysGlyAlaIleTyxThrGluThrGlu----- 365
QY 595 AATCTATTGTCAGCTTTCCTTTTAAAGTTCCTTTGGGAAGTCTTATAGGAGTCTTACTGTT 654
Db 366 -----AspPheSerLeu 369
QY 655 TTAGGAGAGGACATCTGTTGATCTTTCGAGACATACGAGCTTCTCAATGGGAGCT 714
Db 370 LysGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThrGlyGlyAlaLeu 389
QY 715 CTAAGT-----AATAGCGCTGCTGATGCA-----CTGTTTACTATTGAGGCT 756
Db 390 TyxSerLysGlyAsnSerSerLeuThrGlyAsnThrAsnLeuLeuPheSer-----Gly 407
QY 757 TTTAAAGAATTATCTCTTTTCCAAATTCGAATTCATTACTTGCCTGCTGCTGCAACG 816
Db 408 AsnLysAlaThrGlyProSerAsn-----SerSer 417
QY 817 ACTAATAAGGTAGGACGAGCTCCGACGACACATCTACACCGCTCTAATGCTACTATTAT 876
Db 418 AlaAsnGlnGluGlyCysGlyGlyAlaIleLeuAlaPheIleAspSerGlySerValSer 437
QY 877 TCTAAAACAGATCTTTTGTACTCAATATGAGAGTCTCTCTTCTATGATATTAGTC 936
Db 438 AspLysThrGlyLeuSerIleAlaAsnAsnGlnGluValSerLeuThrSerAsnAlaAla 457
QY 937 TCTGAGATGGGGGAGCTATAGATGCTTAAGAGCTTAACGCTTCAAGCAATTTAGCAAGCTT 996
Db 458 ThrValSerGlyGlyAlaIleTyxAlaThrLysCysThrLeuThrGlyAsnGlySerLeu 477
QY 997 TGTGCTCTCCAGAAAATACTGCTCAAGCTGATGGGGAGAGCTTGTCAAGTAGTCACC--- 1053
Db 478 ---ThrPheAspGlyAsnThrAlaGlyThrSerGlyGlyAlaIleTyxThrGluThrGlu 496
QY 1054 AGTTTCTCTGCTATGCTTAACGAGGCTCTCTATTGCTCTTTGAGCAATGTTGAGGAGTA 1113
Db 497 AspPheThrLeuThrGlySerGlyThrValThrPheSerThrAsnThrAlaLysThr 516

Db	1192	SerSer-TyrValIleProLeuHisSerSerLeuArgHisValLeuSerLysValle	1211	Db	3	PheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrLeuSerLeu	22
Qy	2866	-----AGTACGGGTTGG	2879	Qy	979	CAAGGAATTAGCAAGCTTTGTGCTTCCAGAAAATACTGCTCAAGCTGATGGGGAGCT	1038
Db	1211	uProGluLeuProGlyGluThrProLeuValLeuHisGlyGlnValSerTyrGlyArgAs	1231	Db	23	ThrGlyThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGlyGlyAla	42
Qy	2880	GAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGATGTTCTGGGATAA	2939	Qy	1039	TGTCAGTAGTCACACAGTTTCTCTGTATGCGTAAACAGGCTCTCTATTGCTTGTAGCG	1098
Db	1231	nhHisAsnMetThrThrLysLeuAlaAsnAsnThrGlnGlyLysSerAspTrpAspSe	1251	Db	43	lleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPhe	60
Qy	2940	TAACTGTCGCTGGAGAGATTGGAGCGGATTACCGATTGTGATTACTCCATCTAAGCT	2999	Qy	1099	AATGTTGCAGGAGTAAGAGGGGAGGATTGCTGCTGTTCAGATGGCGCAGCAGGAGTG	1158
Db	1251	rHisSerPheAlaValGluValGlyGlySerLeuProValAspLeu	1269	Db	60	-----	60
Qy	3000	CTATTTGAATGAGTTCGCTCTTCTGTCGACGCTGAGTTTCTTATTCGCCGATCATGAATC	3059	Qy	1159	TCATCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTCAGAAATACTCGCGTAGAG	1218
Db	1269	gTyrLeuThrSerTyrSerProTyrValLysLeuGlnValSerValAsnGlnLysG	1289	Db	61	SerAspAsnThrSerSerAspSer	68
Qy	3060	TTTTACAGAGGAGCGATCAAGCTCGGCATTTCAAGAGCGGACATCTCTTAATCTATC	3119	Qy	1219	TTTGATGGGAACGTAGCCCGAGTAGAGGAGGATTTACTCTACGGGAACGTGCTTTC	1278
Db	1289	yPheGlnGluValAlaAspProArgIlePheAspAlaSerHisLeuValAsnValSe	1309	Db	69	-----GlyAlaAlaIlePheThrGluAlaSerValThrIle	80
Qy	3120	AGTTCCTGTTGGAGTGAAGTTTGATGATGTTCTAGTACATCTTAATAATATAGCTT	3179	Qy	1279	CTGAATAATGAAAAACCTTGTCTCTCAACAATGTTGCTTCTCTGTTTACATGCTGCT	1338
Db	1309	rIleProMetGlyLeuThrPheLysHisGluSerAlaLysProProSerAlaLeuLeu	1329	Db	81	SerAsnAsnAlaLysValSerPheIleAspAsn	91
Qy	3180	TATGGCGCTTATCTGTGATGCTTATCGCACATCTCTGCTGCTGAGACAAAGCTCCT	3239	Qy	1339	AAGCAACCAACAGTGGCAGAGCTTCTAATACGAGTAATAATTACGGAGATGGAGGAGCT	1398
Db	1329	uThrLeuGlyTyrAlaValAspAlaTyrArgAspHisProHisCysLeuThrSer	1348	Db	92	-----LysValThrGlyAlaSerSerThrThrGlyAspMet	108
Qy	3240	ATCCCATCAGACATCGACACAGATGCTTTCATTTAGCAACATGAGGTGTGGT	3299	Qy	1399	ATCTTCTGTAGAANTGTCGCAACAGGATCCAAATACCTCTGGATCAGTTTCTCTTGTAT	1458
Db	1348	uThrAsnGlyThrSerTrpSerThrPheAlaThrAsnLeuSerArgGlnAlaPhePhe	1368	Db	109	IleCysAlaTyrLys	123
Qy	3300	TAGAGGATCTATGCTTCTCTTAACAAGTATATAGAAGTATATGCGCATGGAAGATA	3359	Qy	1459	GGAGAGGGAGTAGTTTCTTCTTAGTACCAATGAGTCTGCGGAAAGGGGAGCTATTAT	1518
Db	1368	agluAlaSerGlyHisLeuLysLeuLeuHisGlyLeuAspCysPheAlaSerGlySer	1388	Db	124	GlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrThrAlaGlyGlyAlaIleTyr	143
Qy	3360	TGAGTATCGAGATGCTTCTCGAGGCTATGTTTGAGTGCAGGAGTAGATCGCGTTC	3417	Qy	1519	GCCAAAAGCTCTCGGTCGTCTAATCTGTCGCCCTGTACAAATTTTAAAGGAATATCGCTA	1578
Db	1388	sGluLeuArgSerSerSerArgSerTyrAsnAlaAsnCysGlyThrArgTyrSerPhe	1407	Db	144	ValLysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerValAsn	163
RESULT 24				Qy	1579	-----CATGCTGGAGCGATTATTATTAGGAGAACTCTGGAGAGCTCAGTTTA	1623
H86492	Pmp_3 [imported] - Chlamydophila pneumoniae (strain J138)			Db	164	GlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGluLeuSerLeu	183
C;Species:	Chlamydophila pneumoniae, Chlamydia pneumoniae			Qy	1624	TCTGCTGATATGAGATATATTTTCGATGGGAATCTTAAAGAACACAGCAAGAGAA	1683
C;Date:	02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004			Db	184	SerAlaAspSerGlyAspIleValPheLeuGlyAsn	200
C;Accession:	H86492			Qy	1684	GCTGCCGATGTTAATGCGCTAATCTGCTCTCACAGCCATTTTCGATGGGATCGGAGGG	1743
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie	Nucleic Acids Res. 28, 2311-2314, 2000			Db	201	ThrProGlyThrAsn	215
A;Title:	Comparison of whole genome sequences of chlamydia pneumoniae J138.			Qy	1744	AAAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAG	1803
A;Reference number:	A86491; MUID:20330349; PMID:10871362			Db	216	LysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleTyrPheTyrAspProIleThr	235
A;Accession:	H86492			Qy	1804	ATGGCAACGGAATAACACGAGCCAGCGAGCTTCCAACTTCTAAAAATTAACGATGCT	1863
A;Status:	preliminary			Db	236	ThrGlySerSerThrThr	251
A;Molecule type:	DNA			Qy	1864	GAAGGA	1893
A;Residues:	1-772 <STO>			Db	252	ProAlaAspSerAlaLeuGlnTyrThrGlyAsnIleIlePheThrGlyGluLysLeuSer	271
A;Cross-references:	UNIPROT:Q9RB71; UNIPARC:UPI000000CC33; GB:BA000008; NID:98978389; PI			Qy	1894	-----AATGGAGCAGTACTTTGTACCAAAAATGTTACGATA	1929
A;Experimental source:	strain J138			Db	272	GluThrGluAlaAlaAspSerLysAsnLeuThrSerLysLeuLeuGlnProValThrLeu	291
A;Gene:	pmp_3_2			Qy	1930	GAGCAAGGAGGATTTGTTCTTCGTGAAAGCAAAATATTCAGTGAATTTCTCTAAGTCAG	1989
Alignment Scores:				Qy	919	TTCTATAGTAAATTTAGTCTCTCGAGATGGGAGCTATAGATGCTAAGAGCTTAACGGTT	978
Pred. No.:	2.05e-59	Length:	772	Db	292	SerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThrGlnAlaPheThrGln	311
Score:	979.50	Matches:	261				
Percent Similarity:	46.2%	Conservative:	144				
Best Local Similarity:	29.8%	Mismatches:	323				
Query Match:	12.4%	Indels:	149				
DB:	2	Gaps:	24				
US-10-701-844-1 (1-4435) x H86492 (1-772)							

```
QY 1990 ACAGTGGGAGT---CTGTATATGAAAGCTGGAGTACATGGGATTTTGTAACTCCCAA 2046
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 312 GlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu----- 326
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2047 CCACCACAACAGCCTCTCGCGCTAATCAGTTGATCAGCTTCCAAATCTGCATTTGTCT 2106
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 327 -----ProAlaAspThrSer-----ThrIleAsnAsnLeuValIleAsn 339
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2107 CTTTCTCTTTGTTAGCAACAAGTTCAGAACTCTCTACCAATCTCCAGCGCAA 2166
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 340 IleSerSerIle-----Asp 344
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2167 GATTCCTCATCTCGAGTCATTGGTAGC---ACAACTGCTGGTTCGTGTACAAATAGTGGG 2223
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 345 GlyAlaIleValLeuAlaLeuIleGluThrLysAlaThrSerLysAsnLeuThrLeuSerGly 364
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2224 CCAATCTTTTGGAGATTGGATGATACAGCTTATGATAGTATGATGATGGCTAGGTCT 2283
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 ThrIleThrLeuLeuAspProThrGlyThrPheTyrGluAsnHisSer---LeuArgAsn 383
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2284 AATCAAAAATCAATGCTCTGAAATACAGTTA---GGGACTAAGCCCCAGCTAATGCC 2340
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 384 ProGlnSerTyrAspIleLeuGluLeuLysAlaSerGlyThrValThrSerThrAlaVal 403
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2341 CCATCAGATTTCAGCTCAGGAATCAGATGCCAAGTATGCTATCAGGAAGCTGG--- 2397
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 ThrProAspProIleMetGlyLysPhe---HisTyrGlyTyrGlnGlyThrTrpGly 422
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2398 AAGCTTGCTGGATCTCAATACACAAATAATGCTCTTACTCTGAAGCTACATGG 2457
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 ProIleValTrpGlyThrGlyAlaSerThrThrAlaThrPhe-----AsnTrp 438
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2458 ACTAAAACCTGGTATAATCTCTGGGCTGAGCGAGTAGCTCTCTTGTTCCAAATAGTTTA 2517
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 439 ThrLysThrGlyTyrIleProAsnProGluArgIleGlySerLeuValProAsnSerLeu 458
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2518 TGGGATCCATTTAGATATACGATCTGGCGATTCCAGCAATTCAGCAAGTGTGATGGG 2577
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 459 TrpAsnAlaPheIleAspIleSerSerLeuHisTyrLeuMetGluThrAlaAsnGluGly 478
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2578 CGCTCTTATTGTCGAGGATTAGGGTTCTCTGAGTTTCGAATTTCTTCTATCATCAGCGC 2637
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 479 LeuGlnGlyAspArgAlaPheTrpCysAlaGlyLeuSerAsnProPheHisLysAspSer 498
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2638 GATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTATTCTCTAGGAGCAAACTCC 2697
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 499 ThrLysThrArgArgGlyPheArgHisLeuSerGlyGlyTyrValIleGlyLysAsnLeu 518
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2698 TACTTTGGATCA---TCGATGTTTGGTCTAGCATTTACCGAAGTATTGGTAGACTAAA 2754
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 519 HisThrCysSerAspLysIleLeuSerAlaAlaPheCysGlnLeuPheGlyArgAspArg 538
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2755 GATTATGATGTGTGTTCCAAATCATCTGCTTCATAGGATCCGTTTATCTATCTACC 2814
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 539 AspTyrPheValAlaLysAsnGlnGlyThrValTyrGlyGlyThrLeuTyr-TyrGlnHis 558
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2815 CAACAAGCTTTATGTGGATCTTATTGTTTCGGAGATGCGTTTATCCGTCT----- 2865
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 558 AsnGluThrThrIleSerLeuProCysLysLeuArgProCysSerLeuSerTyrValPr 578
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2866 -----AGCTACGGGTTTGGGAATCAGCATAT 2891
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 578 oThrGluIleProValLeuPheSerGlyAsnLeuSerTyrThrHisThrAspAsnAspLe 598
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2892 GAAACCTCATATACATTTTCAGAGGAGCGCATGTTCTGGGATATAAATCTGCTGCGC 2951
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 598 uLysThrLysTyrThrThrTyrProThrValLysGlySerTrpGlyAsnAspSerPheAl 618
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2952 TCGAGAGATTGGAGCGGATTACCGATTGTGATTACTCCATTAAGCTCTATTTCGAATGA 3011
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 618 aLeuGluPheGlyArgAlaProIleCysLeuAspGluSerAlaLeuPhe---GluGly 637
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 3012 GTTCGGTCTTTCTGTCGAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGA 3071
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 637 nTyrMetProPheMetLysLeuGlnPheValTyrAlaHisGlnGluGlyPheLysGluGly 657
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3072 AGCGCATCAAGCTCGGCATTCAGACGGGACATCTCTAAATCTATCAGTTCCTGTGG 3131
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 657 nGlyThrGluAlaArgGluPheGlySerArgLeuValAsnLeuAlaLeuProIleGly 677
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3132 AGTCAAGTTTGCATCGATGTTCTAGTACATCTCTAAATAATATAGCTTTATGGCGGCTTA 3191
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 677 yIleArgPheAspLysGluSerAspCysGlnAspAlaThrTyrAsnLeuThrLeuGlyTy 697
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3192 TATCTGTGATGCTTATCGCACCATCTCTGGTACTGACACACGCTCTCTATCCCATCA 3251
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 rThrValAspLeuValArgSerAsnProAspCysThrThrThrLeuArgIleSerGlyAs 717
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3252 GACATGGACACACAGATCGCTTTCATTTAGCAAGACATCGAGTGTGTGTAGAGGATCTAT 3311
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 717 pSerTrpLysThrPheGlyThrAsnLeuAlaArgGlnAlaLeuValLeuAlaGlyAs 737
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3312 GTATGCTTCTTCAACAGTAATATAGAGTATATAGCCATGGAAGATATAGATATCGAGA 3371
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 737 nHisPheCysPheAsnSerAsnPheGluAlaPheSerGlnPheSerPheGluLeuArgGly 757
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3372 TGCTTCTCGAGGCTATGTTTGGTGCAGGAAGTAGAGTCCGGTTC 3417
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 ySerSerArgAsnTyrAsnValAspLeuGlyAlaLysTyrGlnPhe 772
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 25
B72131
Polymorphic outer membrane protein g family - Chlamydothila pneumoniae (strain CWL029)
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: B72131
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: B72131
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-922 <ARN>
A:Cross-references: UNIPROT:Q929G5; UNIPARC:UPI000004708A; GB:AE001585; GB:AE001363; NID
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp_1
C:Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 5,998-59 Length: 922
Score: 973.00 Matches: 306
Percent Similarity: 41.5% Conservative: 148
Best Local Similarity: 28.0% Mismatches: 386
Query Match: 12.3% Indels: 253
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x B72131 (1-922)
QY 375 TTGTGTGATGCAACAGCTCTTTCCATAGTCTTCTTCTTCAAT-----GAT 419
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 420 TCTAGCTATTCTTCTCTCTCTTTAAATGCGGGGGGATATGACGAGAAATCATGTGTC 479
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 22 ThrSerLeu-SerAlaThrThrIleSer-----LeuThrPr 33
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 TCAGGAATTTACGATGGGAGACCTTAATCTATCTTCCCTATCTCTTATAGGAGA 539
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 33 oGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyrAsnVal----- 50
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 TCCGAGTGGGACTACTGTGTTTTTTCGCGAGGAGATTAAATTAATAAATCTTGACATTC 599
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 51 -GlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSerAsnValAsnSe 70
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 600 TATTGCGAGCTTTCCCTTTAAGTTGTTTGGGAACCTATTAGGAGGTTTACTGTTTAGG 659
Db 70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGI 88
QY 660 GAGAGGACACTGCTGACTTTTCAGAAACATACGACTTCTTACA-----AATGGGGCAGC 713
Db 88 yAsnHisHisGlyLeuTyPheAsnAsnHisSerSerGlyThrThrLysGluGlyAlaVa 108
QY 714 TCTAAGATAATAGCGCTGCTGATGAGACTGTTTACTATTGAGGGTGTAAAGATTATCTCT 773
Db 108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128
QY 774 TTCCAATTGCAATTCATTACTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
Db 128 e-----IleGI 130
QY 834 GACTCCGACGACAAACATACACCGTCTAATGGTACTATTATTCTTAAACACGACTTTT 893
Db 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTySerLysAsnAlaLeuWe 148
QY 894 GTTACTCAATAATGAGAAGTTCTCATTTCTATAGTAATTTAGTCTCTGGAGATGGGGAGC 953
Db 148 tLeuLeuAsnAsnTy-ValValArgPheGluGlnAsnGlnSerLysThrLysGlyAl 168
QY 954 TATAGATGCTAAGAGCTTAACGTTTCAGGAATTACGAGCTTGTGCTCTTCCAGAA 1013
Db 168 aIleSerGlyAlaAsnValThrIleValGlnAsnTyAspSerValSerPheTyGlnAs 188
QY 1014 TACTGCTCAAGCTGATGGGGGAGCTTGTCAAGTAGTCACCGAGTTCTCTGCTATGGCTAA 1073
Db 188 nAlaAlaThrPheGlyGly-----AlaIleHis 198
QY 1074 CGAGGCTCTTATTCCTTTGACGAATGTCAGGAGTAAGAGGGGAGGATTCCTGTC 1133
Db 198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 212
QY 1134 TGTTCCAGGATGGCAGCAGGAGGTGTCATCTACTTCAACAGAGATCCAGTAGTAAG 1193
Db 212 ----- 212
QY 1194 TTTTCCAGAAATATGCGGTAGAGTTTGATGGAGCTAGTAGCCGAGTAGGAGGAGGAT 1253
Db 213 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 226
QY 1254 TTACTCTCAGCGAAGCTTGCTTCTCTGAATAATGGAAACCTTGTTCTCAACAATGT 1313
Db 226 uTySerAspGlyAspIleAspIleAspGlnAsnAlaTy-ValLeuPheArgGluAsnGI 246
QY 1314 TGCTTCTCTGTTTACATTGCTGTAGCAACCAAGTGGAGAGCTTCTTAATACGAG 1373
Db 246 uAlaLeuThrThrAlaIle----- 252
QY 1374 TAATAATTACGAGATCGAGAGCTATCTTCTGAAGAATGGTGGCAAGCAGGATCCAA 1433
Db 253 -----GlyLysGlyAlaValCysCys-----LeuPr 262
QY 1434 TAACTCTGGATCA-----GTTTCTCTTTGATGGAGAGGAGTAGTGT 1475
Db 262 oThrSerGlySerThrProValProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
QY 1476 CTTTAGTAGCAATGATGCTGCTGGGAAAGGGGAGCTATTATTGCCAAAAGCTCTGGT 1535
Db 282 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyAlaArgLysLeuSerII 302
QY 1536 TGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCT-----AATGA 1580
Db 302 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyAlaAsnSerGlnAsnLe 322
QY 1581 TGTGGAGCGGATTTATTAGGAGAACTCTGGAGAGCTCATGTTTCTGCTGATATGAGGA 1640
Db 322 uGlyGlyAlaIleAlaIleAspThrGlyGlyGluIleSerLeuSerAlaGluLysGlyTh 342
```

```
QY 1641 TATTATTTTCGATGGGAATCTTAAAGAAACAGCCAAAGAGAATGCTGCCGATGTTAATGG 1700
Db 342 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGI 357
QY 1701 CGTAACTGCTGCTCCTCACAGGCATTTCCATGGGATCGGAGGGAATAACGACATTAAG 1760
Db 357 yIleHisLeuLeuGlnAsnAla-----LysPheLeuLysLeuGI 370
QY 1761 AGCTAAAGCAGCGCATCAGATTTCTTTAATGATCCATCGAGATGGCAACGGAATAA 1820
Db 370 nAlaArgAsnGlyTySerIleGluPheTyAspProIle-----Th 384
QY 1821 CCAGCCAGCGCAGTCTTCCAAACTTCTAAATTAAC-----GATGGTGAAGGATA 1871
Db 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
QY 1872 CACGGGATATTGTTTGTCT-----AATCGAAG 1901
Db 404 rThrGlyThrIleLeuPheSerGlyGlyLysSerLeuAlaAsnAspProArgAspPheLy 424
QY 1902 CAGTACTTTGTACCAAAATGTTTACCATAGACAGCAAGAGGATTGTTCTTCGTGAATAAGC 1961
Db 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyLeuValIleLysGluGlyAl 444
QY 1962 AAAATTATCAGTGAATTTCTTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGG 2018
Db 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValIleAspLeuGI 464
QY 2019 GAGTACATGGGATTTGTAACTCCACACCAACACAGCCTCTCCGCGCTAATCAGTT 2078
Db 464 yThr-----LysLeuIleAlaSerLysGluAs 473
QY 2079 GATCAGCTTCCAAATCTGCATTTGCTCTTTCTTTGTTTAGCAACAACATGCACTTAC 2138
Db 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486
QY 2139 GAATCTCTCTACCAATCTCCAGCGCAGAGATTTCTCATCTCGCAGTCATTTGGTAGCAAC 2198
Db 487 -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 499
QY 2199 TGCTGCT-----TCGTTACAAATTAGTGGGCTATCTTT----- 2232
Db 499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
QY 2233 -TTTGAGGATTTGGATGATACAGCTTATGATAGGTATGATTGCTAGGTTCTTAATCAAAA 2291
Db 519 aTyGluAspLeuArg-----MetArgAsnSerGlnTh 530
QY 2292 AATCAATGTCCTGAAATTAACAGTTAGGAGCTAAGCCCGCAGCTAATGCCCATCAGATT 2351
Db 530 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlySerValThrVa 546
QY 2352 GACTCTAGGGAATGAGATG-----CCTAAGTATGGCTATCAAGGAAGCTGGAAGCT 2402
Db 546 lThrAlaGlyAspPheLeuProValSerProHisTyGlyPheGlnGlyAsnTrpLysLe 566
QY 2403 TGGCTGGGATCTTAATACAGCAATAATGCTCTTATCTCTGAAAGCTACATGAGCTAA 2462
Db 566 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspLy 582
QY 2463 AACTGGGTATAATCTCGGCGCTGAGCGAGTAGTCTCTTTGGTTCCAAATAGTTATGGGG 2522
Db 582 sIleAsnTyLysProArgProGlyLysGluGlyAsnLeuValProAsnIleLeuTrpGI 602
QY 2523 ATCCATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGTGAGCGGCTC 2582
Db 602 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGI 622
QY 2583 TTATTGTCGAGGATTTATGGTTTCTCGAGTTTCTCGAATTTCTTCTATCATCACCAGCTGC 2642
Db 622 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerGI 642
QY 2643 TTTAGTTCAGGATATCGGTATATTAGTGGGGTATTCTCTTAGGACAAACTCCTACTT 2702
```

Db 642 uAspAsnIleArgTyrAaGHisAsnSerGlyGlyTyrValLeuSerValAsnAsnGluI 662
 QY 2703 TGGATCATCGATGTTT--GGTCTAGCATTTTACCGAAGTATTTGGTAGATCTTAAGATTA 2759
 Db 662 eThrProIysHisTyrThrSerMetAlaPheSerGlnLeuPheSerAArgAspLysAspTy 682
 QY 2760 TGTAGTGTGTCGTTCCAAATCATCATGCTTGGCATAGGATCCGTTTATCATCTACCCAAACA 2819
 Db 682 rAlaValSerAsnAsnGluTyrArg-----MetTyrLeu----- 693
 QY 2820 AGCTTTATGTGGATCCTATTGTTTC-----GGAGATCGCTTTATCCG 2861
 Db 694 -----GlySerTyrLeuTyrGlnTyrThrThrSerLeuGlyAsnIlePheArgTy 710
 QY 2862 TGCTAGC----- 2868
 Db 710 rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgAArgPheLeuGlnAsnPr 730
 QY 2869 -----TACGGTTTGGGAATCAGCATATGAAACCTC 2900
 Db 730 oLeuMetIlePheHisPheLeuCysAlaTyrGlyHisAlaThrAsnAspMetLysThrAs 750
 QY 2901 ATATACATTTGCAGAGGAGCGATGTTGGTGGATTAATACTCTCTGGCTGGAGAGAT 2960
 Db 750 pTyrAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 770
 QY 2961 TGGAGCGGGATPACCGATGTGATTAATCTCAATCTAAGCTCTATTGAAATGAGTTCGCTCC 3020
 Db 770 eGlyGlySerMetProLeuLeuValPheGlnAsnGlyArgLeuPheGlnGlyAlaIlePr 790
 QY 3021 TTTCGTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGAGGCGATCA 3080
 Db 790 oPheMetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGluThrThrAlaAs 810
 QY 3081 AGCTCGGCAATCAAGAGGGACATCTCTAAATCTATCAGTTCCTGTTGGTAGTGAATT 3140
 Db 810 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830
 QY 3141 TGATCGATGTTCTAGTACACATCTTAATAATATAGCTTTATGCGGCTTATATCTGTGA 3200
 Db 830 eGluLysLeuAlaLeuSerGlnAspValLeuTyrAspPheSerTyrIleProAs 850
 QY 3201 TGCTTATCGCACCATCTCTGGTACTAGACAAACGCTCTCTATCCCATCAAGACATGGAC 3260
 Db 850 pIlePheArgLysAspProSerCysGluAlaAlaLeuValIleSerGlyAspSerTrpLe 870
 QY 3261 AACAGATGCTTTTATAGCAGACATGAGTGTGGTTAGAGATCTATGATGCTTC 3320
 Db 870 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyrHi 890
 QY 3321 TCTAAACAGTAATATAGAAATATATGCGCATGGAAGATATGATATGAGATGCTTCTCG 3380
 Db 890 sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 910
 QY 3381 AGGCTATGTTGAGTGCAGGAAGTAGACTCCGGTTC 3417
 Db 910 gAsnTyrAsnIleAsnCysGlySerLysPheArgPhe 922

RESULT 26
 E86491
 polynorphic outer membrane protein G family [imported] - Chlamydomophila pneumoniae (strai
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: E86491
 R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491, MUID:20330349, PMID:10871362
 A:Accession: E86491
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-922 <STO>

A:Cross-references: UNIPROT:Q929G5; UNIPARC:UPI000004708A; GB:BA000008; NID:98978378; P1
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: pmp 1
 C:Superfamily: Chlamydomophila pneumoniae polynorphic outer membrane protein G

Alignment Scores:
 Pred. No.: 5.99e-59 Length: 922
 Score: 973.00 Matches: 306
 Percent Similarity: 41.5% Conservative: 148
 Best Local Similarity: 28.0% Mismatches: 386
 Query Match: 12.3% Indels: 253
 Gaps: 35

US-10-701-844-1 (1-4435) x E86491 (1-922)

QY 375 TTGTGTGATGCAACAGCTCTTCCATAAGTCTTCTTCTTCAAT-----GAT 419
 Db 5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21
 QY 420 TCTAGCTATTCTTCTGCTCTTTAAATGCGGGGGATATGACAGAGAAATCATGTGTTCC 479
 Db 22 ThrSerLeu-SerAlaThrThrIleSer-----LeuThrPr 33
 QY 480 TCAAGGAATTTACGATGGGAGAGCTTAACTGTATCATTTCCCTATCTACTGTATAGGAGA 539
 Db 33 oGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyrAsnVal----- 50
 QY 540 TCCGAGTGGGACTACTGTTTTTCTGCAGGAGAGTTAAACATTAATAAATCTTGACAAATC 599
 Db 51 -GlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70
 QY 600 TATTGCACTTTGCTTTAAGTGTGTTGGAACTTATTAGGAGTTTTACTGTTTATAGG 659
 Db 70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGl 88
 QY 660 GAGAGGACATCGTTGACTTTTCGAGAGACATACGACTTCTTACA-----AATGGGCGAGC 713
 Db 88 yAsnHisGlyLeuTyrPheAsnAsnIleSerSerGlyThrTyrGlyGlyAlaVa 108
 QY 714 TCTAAGTAATAGCGTGTGATGAGCTGTTTACTTATTGAGGTTTAAAGAAATATCTCTT 773
 Db 108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128
 QY 774 TTCCAATTGCAATTCAATTACTTTCGCGTACTGCTGCTCAACGACTAATAAGGGTAGCCA 833
 Db 128 e-----IleGl 130
 QY 834 GACTCCGAGACACATCTACACCGTCTAATGGTACTATTATTATCTAATAACAGATCTTTT 893
 Db 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTyrSerLysAsnAlaLeuMe 148
 QY 894 GTTACTCAATATGAGAAGTCTCAATCTATATAATTAATTTAGTCTCTGAGAGATGGGGAGC 953
 Db 148 lLeuLeuAsnAsnTyrValValArgPheGluGlnAsnGlnSerLysThrLysGlyAla 168
 QY 954 TATAGATCTAAGAGCTTAAACGTTCAAGGAATAGCAAGCTTTGTGTCTTCCAAAGAAA 1013
 Db 168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyrAspSerValSerPheTyrGlnAs 188
 QY 1014 TACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGGCTAA 1073
 Db 188 nAlaAlaThrPheGlyGly-----AlaIleHisSe 198
 QY 1074 CGAGGCTCTTATGCTTTTGTAGCAATGTTGCAGAGTAAGAGGGGGAGGATGCTGTC 1133
 Db 198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 212
 QY 1134 TGTTCAGGATGGGACGAGGAGGTGTCTATCATCTTACTTCAACAGAAAGATCCAGTAGTAG 1193
 Db 212 ----- 212
 QY 1194 TTTTCCAGAAATACTGCGGTAGATTGATGGGAACGTAGCCCGAGTAGGAGGGGAT 1253


```

Db      213  -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 226
Qy      1254 TTACTCTACGGGAAGCTGCTTCTCTGAATAATGGAAAAACCTGTTTCTCAACAATGT 1313
Db      226  uTyRSerAspGlyAspIleAspIleAspGlnAsnAlaTyValLeuPheArgGluAsnG1 246
Qy      1314 TGTCTTCTCTGTTTACATTGCTGCTAAGCAACCAAGTGGACAGGCTTCTAATACGAG 1373
Db      246  uAlaLeuThrThrAlaIle----- 252
Qy      1374 TAATAATTACGGAGATGGAGGAGCTATCTTCTGAAGAATGTGCGCAACGAGATCCAA 1433
Db      253  -----GlyLysGlyGlyAlaValCysCys-----LeuPr 262
Qy      1434 TAACCTCGATCA-----GTTCTCTTGATGGAGAGGAGTAGTATTT 1475
Db      262  oThrSerGlySerSerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
Qy      1476 CTTTAGTAGCAATGTAGCTGCTGGGAAGGGGAGCTATTATGCCCCAAAAGCTCTCGGT 1535
Db      282  lPheGluArgAsnHisSerIleMetGlyGlyGlyAlaIleTyAlaArgLysLeuSerIl 302
Qy      1536 TGTCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCT-----AATGA 1580
Db      302  eSerSerGlyGlyProThrLeuPheIleAsnAlaIleSerTyAlaAsnSerGlnAsnLe 322
Qy      1581 TGTGTGAGCGCAATTTATTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTTATGGAGA 1640
Db      322  uGlyGlyAlaIleAlaIleAspThrGlyGlyGlyGluIleSerLeuSerAlaGluLysGlyTh 342
Qy      1641 TATTATTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAATGCTGCCGATGTTTATGCG 1700
Db      342  rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnG1 357
Qy      1701 CGTAACTGTCTCTCACAAAGCCATTTTCGATGGGATCGGGAGGAAATAACGACATTAAG 1760
Db      357  yIleHisLeuGlnAsnAla-----LysPheLeuLysLeuG1 370
Qy      1761 AGCTAAAGCAGGCGATCAGATTCTCTTAAATGATCCCATGCGAGATGGCAACCGAAATAA 1820
Db      370  nAlaArgAsnGlyTyRSerIleGluPheTyAspProIle-----Th 384
Qy      1821 CCAGCCAGCGCAGCTTCCAAACTCTTAAATAATTAAC-----GATGCTGAAGATA 1871
Db      384  rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
Qy      1872 CACAGGCGATATGTTTTCGT-----AATGGAAG 1901
Db      404  rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheLy 424
Qy      1902 CAGTACTTTGTACCAAAATGTTACGATAGACAGGAAGGATTCTCTCGTGAAGAGGC 1961
Db      424  sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyRLeuValIleLysGluGlyAl 444
Qy      1962 AAAATTATCAGTGAATCTCTTAAGTCACAGAGTGGGAGT---CTGTATATGGAAGCTGG 2018
Db      444  aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuG1 464
Qy      2019 GAGTACATGGGATTTTGTAATCTCCACACACCAACGAGCTCTCTCGCGCTAATCAGTT 2078
Db      464  yThr-----LysLeuIleAlaSerLysGluAs 473
Qy      2079 GATCAGCGCTTCCAAATCTGATTTGTCTCTTCTTCTTTGTAGCAAAATGCAAGTTAC 2138
Db      473  pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486
Qy      2139 GAATCTCTCTACCAATCTCCAGCGCAAGATTCTCATCTCCTCGAGTATGTTAGACAAAC 2198
Db      487  -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 499
Qy      2199 TGTCTGGT-----TCTGTACAAATTAGTGGGCTATCTTT----- 2232

```

```

499  rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
2233  -TTTCAGAGATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTTCTTAACAAA 2291
519  aTyRGluAspLeuArg-----MetArgAsnSerGlnTh 530
2292  AATCAATGTCCTGAATAATTACAGTTTAGGACTAAGCCCCAGCTAATAGCCCCATCAGATT 2351
530  rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlyGlySerValThrVa 546
2352  GACTCTAGGAATGATAGTG-----CCTAAGTATGGCTATCAAGGAAGCTGGAAGCT 2402
546  lThrAlaGlyAspPheLeuProValSerProHisTyRGlyPheGlnGlyAsnTrpLysLe 566
2403  TCGGTGGGATCTCTAATACAGCAATAATGGTCTCTTATCTCTGAAAGCTACATGAGCTAA 2462
566  uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspLy 582
2463  AACTGGGTATATCTCGGCCCTGAGCGAGTACGCTCTTCTTGGTTCCAAATAGTTTATGGGG 2522
582  sIleAsnTyRlysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpG1 602
2523  ATCCATTTTATGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGATGGGCGCTC 2582
602  yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuG1 622
2583  TTATTGTCGAGGATATGGGTTTCTGGAGTTTTCGAATTTCTTCTATCATGACCGCATGC 2642
622  nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPhePheHisValSerAlaSerG1 642
2643  TTTAGTTCAGGATATCGGTATATTAGTGGGGTTTATCTTCTTAGGACCAACTCTCTACT 2702
642  uAspAsnIleAArgTyArgHisAsnSerGlyTyRValLeuSerValAsnAsnGluI1 662
2703  TGGATCATCGATGTTT---GGTCTAGCATTTACCGAAGTATTGTTGGTAGATCTTAAGATT 2759
662  eThrProLysHisTyThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 682
2760  TGTAGTGTGCTGCTTCCAATCATCATGCTTGCATAGGATCGGTTTATCTATCTACCCAACA 2819
682  rAlaValSerAsnAsnGluTyArg-----MetTyLeu----- 693
2820  AGCTTTTATGTCGATCTCTATTGTTTC-----GGAGATGCTGTTTATCCG 2861
694  -----GlySerTyLeuTyRGlnTyRThrThrSerLeuGlyAsnIlePheArgTy 710
2862  TGCTAGC----- 2868
710  rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr 730
2869  -----TACGGGTTTGGGAATCAGCATATGAAAACCTC 2900
730  oLeuMetIlePheHisPheLeuCysAlaTyRGlyHisAlaThrAsnAspMetLysThrAs 750
2901  ATATACATTTCAGAGAGAGAGCGATGCTGCTCGGATAATAACTGTCTGCTCGAGAGAT 2960
750  pTyAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 770
2961  TGGAGCGGGATTCACGATTGTGATTACTCCATAGCTCTATTGTAAGTGGTGGCTGCC 3020
770  sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 790
3021  TTTTCGTCGACAGTGGATTTTCTTATCGGATCATGATCTTTTACGAGGAGGCGCATCA 3080
790  oPheMetLysLeuGlnLeuValTyRAlaTyRGlnGlyAspPheLysGluThrThrAlaAs 810
3081  AGCTCGGGCATTCAGAGCGGCATCTCTCAATCTATCATCTCTGCTGCTGAGTGAAGTT 3140
810  pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830
3141  TGATCGATGTTCTAGTACACATCTCTAATAAATATAGCTTTTATGGCGGCTTATATCTGTA 3200
830  eGluLysLeuAlaLeuSerGlnAspValLeuTyRAspPheSerPheSerTyRileProAs 850

```


QY 3201 TGCTTATCGCCACCATCTCTGGTACTGAGACAAAGCGTCTCTATCCCATCAAGAGACATGGAC 3260
Db : : : : :
850 pillePheArgLysAspProSerCysGluAlaLeuValIleSerGlyAspSerTrpLe 870
QY 3261 AACAGATGCCCTTTCATTTAGCAGACAGGAGTGTGGTTAGAGATCTATGTATGCTTC 3320
Db : : : : :
870 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyRHi 890
QY 3321 TCTAACAGTAATATAGTAAGTATAGGCATGAGATATGATGATCGATGCTTCTTCG 3380
Db : : : : :
890 sPheAsnAspTyRThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 910
QY 3381 AGGCTATGGTTGAGTCGAGGAGTAGAGTCGGTTC 3417
Db : : : : :
910 gAenTyRAsnIleAsnCysGlySerLysPheArgPhe 922

RESULT 27
F81539
polymorphic membrane protein G family CP0770 [imported] - Chlamydomophila pneumoniae (stra
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
A:Accession: F81539
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
J. C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: F81539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-922 <REA>
A:Cross-references: UNIPARC:UP100016560D; GB:AE002237; GB:AE002161; NID:g7189684; PIDN:
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0770
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Alignment Scores:
Pred. No.: 7, 02e-59 Length: 922
Score: 972.00 Matches: 306
Percent Similarity: 41.4% Conservatives: 147
Best Local Similarity: 28.0% Mismatches: 387
Query Match: 12.3% Indels: 253
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x F81539 (1-922)

QY 375 TTGTGTGATGCAACGCTCTTCCATAGTCTTCTTCAAT-----GAT 419
Db : : : : :
5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21
QY 420 TCTAGCTTATCTTCTGCTCTTTAAATGGGGGGGATATGCGAGAAATCATGGTTC 479
Db : : : : :
22 ThrSerLeu-SerAlaThrIleSer-----LeuThrPr 33
QY 480 TCAAGAAATTTACGATGGGAGACGCTAACTGTATCTTCCCTATCTGTTATAGGAGA 539
Db : : : : :
33 OgluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyRAsnVal----- 50
QY 540 TCCGAGTGGGACTACTGTTTTTTCGAGAGAGTAACTTAAATAATCTTGAATTC 599
Db : : : : :
51 -GlnAlaGlyAspValTyR-SerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70
QY 600 TATTGCGAGTTCCTTTAAAGTGTGTTGGGAACCTATTAGGAGCTTTTACTGTTTAGG 659
Db : : : : :
70 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGI 88
QY 660 GAGAGGACACTCGTTGACTTTTCGAGAACATAGGACTTCTACA-----AATGGGGCAGC 713
Db : : : : :
88 yAenHisHisGlyLeuTyR-PheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 108
QY 714 TCTAAGTAATAGCGCTGCTGATGGAGCTGTTTACTATTGAGGGTTTTTAAGAAATATCCTT 773

108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128
QY : : : : :
774 TTCCAATTTGCAATTCATTTACTTGGCGTACTGCTGCTGCAACGACTAATAGGGTAGCCA 833
Db : : : : :
128 e-----11eGI 130
QY 834 GACTCCGACGACAACTACTACCGTCTAATGCTACTATTATTCTTCTTCTTCTTCTTCTTCT 893
Db : : : : :
130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTyRSerLysAsnAlaLeuMe 148
QY 894 GTTACTCAATTAATCAGAAGTCTCATTTCTATAGTAATTTAGTCTCTCGAGATGGGGAGC 953
Db : : : : :
148 tLeuLeuAsnAsnTyRValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 168
QY 954 TATAGATGCTAAGAGCTTAAAGGTTCAAGGAATTTAGCAAGCTTTGTCTCTTCTTCTTCTTCT 1013
Db : : : : :
168 alleSerGlyAlaAsnValThrIleValGlyAsnTyRAspSerValSerPheTyRGlndAs 188
QY 1014 TACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTTTCTCTCTATGCTTAA 1073
Db : : : : :
188 nAlaAlaThrPheGlyGly-----AlaIleHisSe 198
QY 1074 CGAGGCTCTTATTCCTTTGTAGCAATGTTGCGAGTAGTAAGAGGGGAGGATGCTGTC 1133
Db : : : : :
198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 212
QY 1134 TGTTGAGATGGGAGGAGGAGTGTCTATCTTCAACAGAGATCCAGTAGTAAG 1193
Db : : : : :
212 ----- 212
QY 1194 TTTTCCAGAAATACTCGGTAGATTGTGATGGAAGCTAGCCGAGTAGGAGGAGAT 1253
Db : : : : :
213 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 226
QY 1254 TTACTCTTACGGGAACGTTGCTTCTGTAATTAATGGAACCTTGTCTCTCAACATGT 1313
Db : : : : :
226 uTyR-SerAspGlyAspIleAspIleAspGlnAsnAlaTyRValLeuPheArgGluAsnGI 246
QY 1314 TGCTTCTCTGTTTACATTTGCTGTAAGCAACCAAGTGGACAGGCTTCTAATACGAG 1373
Db : : : : :
246 uAlaLeuThrThrAlaIle----- 252
QY 1374 TAATAATTAACGAGATGAGAGGAGTCTTCTGTAAGATGCTGCGAGAGGATCCAA 1433
Db : : : : :
253 -----GlyLysGlyGlyAlaValCysCys-----LeuPr 262
QY 1434 TAACCTCGATCA-----GTTCTCTTTCATGAGAGGAGTAGTCTT 1475
Db : : : : :
262 oThrSerGlySerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
QY 1476 CTTTAGTAGCAATCTAGCTGCTGGGAAAGGGGAGCTATTATTCGCAAAAGCTCTCGGT 1535
Db : : : : :
282 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyRAlaArgLysLeuSerIl 302
QY 1536 TGCTAATCTGCGCTCTGTAACATTTTAAAGAAATATCGCT-----AATGA 1580
Db : : : : :
302 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyRAlaAsnSerGlnAsnLe 322
QY 1591 TGGTGGAGCGATTTATTTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGA 1640
Db : : : : :
322 uGlyGlyAlaIleAlaIleAspThrGlyGlyIleSerLeuSerAlaGluLysGlyTh 342
QY 1641 TATTATTTCGATGGGAATCTTAAAGAAACAGCAGCAAGAGATGCTCGGATGTTAATGG 1700
Db : : : : :
342 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGI 357
QY 1701 CGTAACTGTGCTCTCAAGACCATTTTCATGGGATCGGAGGGGAAATTAACGACATTAAG 1760
Db : : : : :
357 yIleHisLeuLeuGlnAsnAla-----LysPheLeuLysLeuGI 370
QY 1761 AGCTAAAGCAGGCGATCAGATTTCTTCTTAAATGATCCCATCGAGATGGCAACGGAATAA 1820
Db : : : : :
1820

Db 370 nAlaAsgAsnGlyTyrSerIleGluPheTyrAspProIle-----Th 384
Qy 1821 CCGCCAGCGCAGTCTTCCAACTTCTAAATAATAC-----GATGGTGAAGATA 1871
Db 384 rSerGluAlaAAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
Qy 1872 CACAGGGGATATTGTTTTTGGCT-----AATGGAAG 1901
Db 404 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheLy 424
Qy 1902 CAGTACTTTGTACCAAAATGTTACGATAGACAGGAAGGATGTTCTTCGTGAAGGC 1961
Db 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyrLeuValIleLysGluGlyAl 444
Qy 1962 AAATATCATCGTAATCTCTAGTCAGACAGTGGAGT---CTGTATATGAACTCG 2018
Db 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGl 464
Qy 2019 GAGTACATGGGATTTGTAACTCCACCAACACCAACAGCTCTCGCCCTCAATCAGTT 2078
Db 464 yThr-----LysLeuIleAlaSerLysGluAs 473
Qy 2079 GATCAGCGTTTCCAATTCGATTTGTCTCTTTCTTTCTTTCTTTAGCAAAATCAGTTAC 2138
Db 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486
Qy 2139 GAATCTCTTACCAATCTCCAGCGCAGAGATTCTCATCTGCGAGTCATTTGGTAGCACAC 2198
Db 487 -----SerSerSerSerThrAlaAlaValIleLysAlaAsnTh 499
Qy 2199 TCGTGGT-----TCTGTCAATAGTGGCCCTATCTTT----- 2232
Db 499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
Qy 2233 -TTTGAGGATTTGCGATGATACAGCTTATGATAGTATGATGCTAGGTTCTTAATCAAAA 2291
Db 519 aTyrGluAspLeuArg-----MetArgAsnSerGlnTh 530
Qy 2292 AATCAATGCTCTGMAATTACAGTTAGGAGTAAAGCCCGAGCTAATGCCCATCAGATTT 2351
Db 530 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlyGlySerValThrVa 546
Qy 2352 GACTCTAGGGAATCAGATG-----CCTAAGTATGGCTATCAAGAAAGCTGGAAGCT 2402
Db 546 lThrAlaGlyAspPheLeuProValSerProHisTyrGlyPheGlnGlyAsnIlePylsLe 566
Qy 2403 TGGCTGGGATCCTAAATACAGCAATATGCTCTTATCTACTCTGAAGCTACATGGACTAA 2462
Db 566 uAlaTrpThrGlyThrGlyAsnLysValGlyGluPhe-----PheTrpAspLy 582
Qy 2463 AACTGGGTATATCTCGGGCTGAGCAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGG 2522
Db 582 sIleAsnTyrLysProArgProGluLysGluGlyAsnLeuValProAsnIleLeuTrpGl 602
Qy 2523 ATCCATTTAGATATAGCTGCGCATTCAGCAATTCAGCAATTCAGCAAGTGTGGATGGCGCTC 2582
Db 602 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuGl 622
Qy 2583 TTATTGTCGAGGATATGAGGTTTCTGGAGTTTCCGAATTTCTTCATCATCAGCCGATGC 2642
Db 622 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAsnPheHisValSerAlaSerGl 642
Qy 2643 TTTAGGTCAGGAGTATCGGTATATATAGTGGGGTTATTCCTTAGGAGCAAACTCTACTT 2702
Db 642 uAspAsnIleArgTyrArgHisAsnSerGlyGlyTyrValLeuSerValAsnAsnGluIle 662
Qy 2703 TGGATCATCGATGTT---GCTCTAGCATTTACCAAGTATTTGGTAGATCTAAAGATTA 2759
Db 662 eThrProLysHisTyrThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 682
Qy 2760 TGTAAGTGTGCTTCCATCATCATGCTTGCATAGGATCCCTTTATCTATCTATCCCAACA 2819
Db 682 rAlaValSerAsnAsnGluTyrArg-----MetTyrLeu----- 693

Qy 2820 AGCTTTATGTGATCCTATTGTTCT-----GGAGATCGGTTTATCCG 2861
Db 694 -----GlySerTyrLeuTyrGlnTyrThrThrSerLeuGlyAsnIlePheArgTy 710
Qy 2862 TGCTAGC----- 2868
Db 710 rAlaSerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnPr 730
Qy 2869 -----TACGGGTTTGGGAATCAGCATATGAAACCTC 2900
Db 730 oLeuMetIlePheHisPheLeuCyAlaTyrGlyHisAlaThrAsnAspMetLysThrAs 750
Qy 2901 ATATACATTTGCAGAGGAGGAGCATGTTCTGTTGGGATAATAACTGTCTCGCTGGAGAT 2960
Db 750 pTyrAlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCy 770
Qy 2961 TGGAGCGGATTCACGATTGATTAATCTCATCTAGCTCTATTTGAATAGTTGGTGC 3020
Db 770 sGlyGlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIlePr 790
Qy 3021 TTTCTGCAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGAGGCGCATCA 3080
Db 790 oPheMetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGlnThrThrAlaAs 810
Qy 3081 AGCTCGGCAATTCAGAGCGGACATCTCCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTT 3140
Db 810 pGlyArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 830
Qy 3141 TGATCGATGTTCTAGTACATCTCTTAATAATATAGCTTTATGCGGCTTATATCTGTGA 3200
Db 830 eGluLysLeuAlaLeu**GlnAspValLeuTyrAspPheSerPheSerTyrIleProAs 850
Qy 3201 TGTCTTATCCACCATTCTGTTACTGAGACACAGCTCCTATCCCATCAAGACATGGAC 3260
Db 850 pIlePheArgLysAspProSerCysGluAlaLeuValIleSerGlyAspSerTrpLe 870
Qy 3261 AACAGATGCTTTCATTTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCTTC 3320
Db 870 uValProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyrHi 890
Qy 3321 TCTAACAGTAATATAGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 3380
Db 890 sPheAsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 910
Qy 3381 AGCTATGTTTGCAGTGCAGGAGTACAGTCCGGTTC 3417
Db 910 gAsnTyrAsnIleAsnCysGlySerLysPheArgPhe 922

RESULT 28

E72130

polymorphic membrane protein G family CP0761 [imported] - Chlamydothila pneumoniae (str C; Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C; Accession: E72130; G81541
R; Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21: 385-389, 1999

A; Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A; Reference number: A72000; MUID:99206606; PMID:10192388
A; Accession: E72130
A; Molecule type: DNA

A; Residues: 1-841 <ARN>
A; Cross-references: UNIPROT:Q9Z3A1; UNIPARC:UPI00000470B5; GB:AE001586; GB:AE001363; N1
A; Experimental source: strain CWL029
R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey, R.; C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A; Reference number: A81500; MUID:20150255; PMID:10684935
A; Accession: G81541
A; Molecule type: DNA
A; Residues: 1-841 <REA>
A; Cross-references: UNIPARC:UPI00000470B5; GB:AE002235; GB:AE002161; NID:g7189672; PID:

A: Experimental source: strain AR39, HL cells

C: Genetics:

A: Gene: pmp_2; CP0761

C: Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Alignment Scores:

Pred. No.:	1.19e-56	Length:	841
Score:	939.50	Matches:	289
Percent Similarity:	39.5%	Conservative:	134
Best Local Similarity:	27.0%	Mismatches:	358
Query Match:	11.9%	Indels:	291
DB:	2	Gaps:	30

US-10-701-844-1 (1-4435) x E72130 (1-841)

QY	382	ATGCAAGCTCTTCCATAAGTTCTTCTTCAATGATTTAGCTTATTTCTTGCTCT	441
DB	1	MetLysIleProLeuArgPheLeuLeuIleSerLeuValProThrLeuSerMetSer	20
QY	442	TTAAATGGGGGGATATGCAGCAATATCATGGTCTCTCAAGGAATTTACGATGGGAG	501
DB	21	LeuLeuGlyAlaAlaThrThrGluGluLeuSerAlaSerAsnSer---	39
QY	502	ACGTTAACTGTATCATTTCCCTATCTCTTATAGGAGATCCGAGTGGGACTCTGTTT	561
DB	40	ThrSerThrThrSerPheSerSerLysThrSerSerAlaThrAspGlyThrAsnTyrVal	59
QY	562	TCTGAGGAGAGTTAAACATTTAAATAATCTTGACATTTCTATTGCAGCTTTGCTTTAAGT	621
DB	60	PheLysAspSerValValIleGluLeuValProLysThrGlyThrGlnSerThrSer	79
QY	622	TGTTTGGGAAC-----TTATTAGGAGTTTACTGTGTTTATAGGAGAGGACACTCGTTG	675
DB	80	CysPheLysAsnAspAlaAlaAlaGlyAspLeuAsnPheLeuGlyGlyPheSerPhe	99
QY	676	ACTTTCGAGAACATA---CGGACTTCTACAAATGGGAGCTCTAAGTAATAGCCGCTCT	732
DB	100	ThrPheSerAsnIleAspAlaThrThrAlaSerGlyAlaAlaIleGlySerGluAlaAla	119
QY	733	GATGACTGTTTACTATTGAGGTTTAAAGAAATTTCTTTTCCAAATTTGCAATTCATTA	792
DB	120	AsnLysThrValThrLeuSerGlyPheSerAlaLeuSerPhe-----	133
QY	793	CTTGCGTACTGCTGCTGCAAGCACTAATAAGGTAGCCAGACTCCGACGACACACTCT	852
DB	134	-----LeuLysSerProAlaSerThrVal	141
QY	853	ACACCGTCTAATGGTACTATTATTCTTAAACACAGATCTTTTCTTACTCAATAATCAGAAG	912
DB	142	ThrAsnGlyLeuGlyAlaAlaIleAsnValLysGlyAsnLeuSerLeuLeuAspAsnLys	161
QY	913	TTCTCATTTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGAT---GCTAAGAGC	969
DB	162	ValLeuIleGlnAspAsnPheSerThrGlyAspGlyAlaIleAsnCysAlaGlySer	181
QY	970	TTAAGGTTTCAAGGAATTAGCAAGCTTTGTGCTCTCCAAAGAAATACTGCTCAAGCTGAT	1029
DB	182	LeuLysIle-----	184
QY	1030	GGGGAGCTGTGCAAGTAGTACCAGTTTCTCTGCTATGCTTAAAGAGGCTCTATTGGC	1089
DB	185	-----AlaAsnAsnLysSerLeuSer	191
QY	1090	TTTGTAGCGAATTTGTCAGGAGTAAAGGGGGAGGATTTGCTGCTTCCAGGATGGCAG	1149
DB	192	PheIleGlyAsnSerSerThrArgGlyGlyAlaIleHisThr-----	206
QY	1150	CAGGGAGTGTCTATCTACTTCAACAGAAGATCCAGTAGTAGTATTTTCCAGAAATACT	1209
DB	206	-----	206
QY	1210	GGGTAGAGTTTGATGGGAACGTAGCCCGAGTAGGAGGAGGATTTTACTCTACGGGAC	1269

DB	207	-----LysAsn	208
QY	1270	GTTGCTTTCGTAATAATGGAACCTTGTCTTCTCAACATGTTGCTTCTCCTGTTTAC	1329
DB	209	LeuThrLeuSerSerGlyGlyLeuThrLeuPheGlnGlyAsnThrAla-	224
QY	1330	ATTGCTGCTAAGCAACCAAGTGGACAGCTTCTTAATACGAGTAATAATTACGAGAT	1389
DB	225	-----ProThr	226
QY	1390	GGAGGAGCTATCTTCTGTAAGAATGTGTGCGCAAGCAGGATCCAATAACTCTGATCAGTT	1449
DB	226	-----	226
QY	1450	TCCTTTGATGAGAGGAGTAGTTTCTTCTTAGTAGCAATGTAGCTGCTGGGAAAGGGGA	1509
DB	227	-----AlaAlaGlyLys	230
QY	1510	GCTATTTATGCCAAAAGCTCTCGTTGCTTAACGTGCGCCCTGTACAATTTTAAAGGAAT	1569
DB	230	-----	230
QY	1570	ATCGCTAATGATGGTGGAGCGATTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCT	1629
DB	231	-----GlyGlyAlaIleAlaAlaAspSerGlyThrLeuSerIleSerGly	246
QY	1630	GATTATGAGATATTTTTCGATGGGAATCTTAAAGAACAGCAAGAAAGATGCTGCC	1689
DB	247	AspSerGlyAspIleIlePheGluGly-----AsnThrIle	258
QY	1690	GATGTTAATGCGTAACTGTGCTCTCAAGCCATTTTCGATGGATGGGAGGAAATA	1749
DB	259	GlyAlaThrGly---ThrValSerHisSerAlaIleAspLeuGlyThrSerAlaLysIle	277
QY	1750	ACGACATTAAAGAGCTAAAGCAGGAGCATCAGATTCTCTTATATGATCCCATCGAGTGGCA	1809
DB	278	ThrAlaLeuAlaAlaGlnGlyHisThrIleThrPheThrAspProIleThrVal---	296
QY	1810	AACGAAATAACAGCCAGCCAGCTCTTCCAACTTCTTAAATAATTAAC-----	1857
DB	297	ThrGlySerThrSerValAlaAspAla-----LeuAsnIleAsnSerProAspThr	313
QY	1858	---GATGTTGAAGGATACAGGGGATATGCTTTTGTCT-----	1893
DB	314	GlyAspAsnLysGluThrGlyThrIleValPheSerGlyGlyLysLeuThrGluAla	333
QY	1894	-----AATGGAAGCAGTACTTTGTACCAAAATGTTTACGATAGAGCAA	1935
DB	334	GluAlaLysAspGluLysAsnArgThrSerLysLeuLeuGlnAsnValAlaPheLysAsn	353
QY	1936	GGAAGGATTTGTTCTGTAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGT	1995
DB	354	GlyThrValValLeuLysGlyAspValValLeuSerAlaAsnGlyPheSerGlnAspAla	373
QY	1996	GGGAGT---CTGTATATGGAAGCTGGGAGTATCGGATTTTGTAACTCCACCAACCA	2052
DB	374	AsnSerLysLeuIleMetAspLeuGlyThrSerLeu-----	385
QY	2053	CAACAGCCTCTCGCGCTAATCAGTTGATCAGCTTTTCCAAATCTGCAATTTGTCTTTCT	2112
DB	386	-----ValAlaAsnThrGluSerIleGluLeuThrAsnLeuGluIleAsnIleAsp	402
QY	2113	TCTTTGTTAGCAACAATGCGATTCAGAACTCTCTTACCAATCTCCAGCGCAGATTCT	2172
DB	403	SerLeuArgAsnGlyLysLysIle-----	410
QY	2173	CATCTCGAGTCATTTGTTAGGCAACTGCTGTTCTGTTTACAAATTTAGTGGGCTATCTTT	2232
DB	411	-----LysLeuSerAlaAlaThrAlaGlnLysAspIleArgIleAspArgProVal	428
QY	2233	TTTGAGGATTTGATGATACACAGCTTATGATAGG-----TAT	2268
DB	429	LeuAlaIleSerAspGluSerPheThrGlnAsnGlyPheLeuAsnGluAspHisSerThr	448

```
QY 2269 GATTGGCTAGGTTCTTAATCAAAAATCAATGTCCTGABATTACAGTTAGGAGTAAAGCCC 2328
Db 449 Asp -----GlyIleLeuGluLeuAspAlaGlyLysAspIle 460
QY 2329 CCAGCTAAATGCCCATCAGATTGACTCTAGGGAATGAGATGCTTAAGTATGCTGCTATCAA 2388
Db 461 ValIleSerAlaAspSerArgSerIleAspAlaValGlnSerPro---TyrGlyTyrGln 479
QY 2389 GGAGCTGGAGCTTCGCTGGGATCCTTAATACAGCAATAATAGTCTTATCTATCTGAAA 2448
Db 480 GlyLeuTrpThrIleAsnTrpSerThrAspAspLys-----Lys 492
QY 2449 GCTTACA-----TGGACTAAACTGGGTATATCTCTGGCGCTGAGCGAGTAGCTTCTTTG 2502
Db 493 AlaThrValSerTrpAlaLysGlnSerPheAsnProThrAlaGluGlnGluAlaProLeu 512
QY 2503 GTTCCAAATAGTTTATGGGATCCATTTAGATATACGATTCGCGCATTCAGCAATTCAA 2562
Db 513 ValProAsnLeuLeuTrpGlySerPheIleAspValArgSerPheGlnAsnPheIleGlu 532
QY 2563 GCAAGTGTGATGGCGCTCTTATGTCGAGGATTAATGGTTCTCGAGTTTCGAATTC 2622
Db 533 LeuGlyThrGluGlyAlaProTyrGluLysArgPheTrpValAlaGlyIleSerAsnVal 552
QY 2623 TTCATCATGACCGCGATGCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATTC 2682
Db 553 LeuHisArgSerGlyArgGluAsnGlnArgLysPheArgHisValSerGlyIleAlaVal 572
QY 2683 TTAGGAGCAACTCTACTTT---GGATCATCGATGTTTGGTCTAGCATTTACCGAAGTA 2739
Db 573 ValGlyAlaSerThrArgMetProGlyGlyAspThrLeuSerLeuGlyPheAlaGlnLeu 592
QY 2740 TTTGGTAGATCTAAAGATTATGATGTGTGTCCTCAATCATCATGCTTCGATAGATCC 2799
Db 593 PheAlaArgAspLysAspTyrPheMetAsnThrAsnPheAlaLysThrTyrAlaGlySer 612
QY 2800 GTTATCTATCTACCCACAGCTTTATGTGGA-----TCCTATTGTTTCGAGATGCG 2853
Db 613 LeuArgLeuGlnHisAspAlaSerLeuTyrSerValValSerIleLeuLeuGlyGluGly 632
QY 2854 TTTATCGT----- 2862
Db 633 GlyLeuArgGluLeuLeuLeuProTyrValSerLysThrLeuProCysSerPheTyrGly 652
QY 2863 ---GCTAGCTACGGTTTGGGAATCAGCATATGAAATCACTCATAT----- 2904
Db 653 GlnLeuSerTyrGlyHisThrAspHisArgMetLysThrGluSerLeuProProPro 672
QY 2905 ---ACATTCGAGAGGAGCGATGTCGTTGGGATAATACTGTCTGGCTGGAGAGATT 2961
Db 673 ProThrLeuSerThrAsp---HisThrSerTrpGlyGlyTyrValTrpAlaGlyGluLeu 691
QY 2962 GGAGCGGATACCGATTGATGTATCTCCATCAAGCTCTATTGGAATGATGTCGCT 3021
Db 692 GlyThrArgValAlaValGluAsnThrSerGlyArgGlyPhePheGlnGluTyrThrPro 711
QY 3022 TTCGTGCAAGCTGAGTTCTTTATGCGCATCATCAATCTTTTACAGAGGAAGCGATCAA 3081
Db 712 PheValLysValGlnAlaValTyrAlaArgGlnAspSerPheValGluLeuGlyAlaIle 731
QY 3082 GCTCGGCATTCAGAGCGGACATCTCTAAATCATCATCAGTTCCTGTGGAGTGAAGTTT 3141
Db 732 SerArgAspPheSerAspSerHisLeuTyrAsnLeuAlaIleProLeuGlyIleLysLeu 751
QY 3142 GATCGATGTTCTAGTACATCTCAATTAATATAGCTTTATGGCGGCTTATATCTGAT 3201
Db 752 GluLysArgPheAla-----GluGlnTyrTyrHisValValAlaMetTyrSerProAsp 769
QY 3202 GCTTATCGCACATCTCTGGTACTGAGACACGCTCTCTATCCCATCAAGAGACATGGACA 3261
Db 770 ValCysArgSerAsnProLysCysThrThrThrLeuLeuSerAsnGlnGlySerTrpLys 789
```

```
QY 3262 ACAGATGCGCTTTCAATTAGCAAGACATGGAGTGTGGTTAGGAGATCTATGATGCTTCT 3321
Db 790 ThrLysGlySerAsnLeuAlaArgGlnAlaGlyIleValGlnAlaSerGlyPheArgSer 809
QY 3322 CTAACAAGTATATAGAATGATATGCGCATGCAAGATATGATGATCGAGATCTCTCCA 3381
Db 810 LeuGlyAlaAlaAlaGluLeuPheGlyAsnPheGlyPheGluTrpArgGlySerSerArg 829
QY 3382 GCGTATGTTTTCAGTGTGAGAGTAGATGTCGGTTC 3417
Db 830 SerTyrAsnValAspAlaGlySerLysIleLysPhe 841
RESULT 29
F81721
polymorphic membrane protein G family TC0267 [imported] - Chlamydia muridarum (strain N1)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81721
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: F81721
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-867 <DET>
A:Cross-references: UNIPROT:Q9PL41; UNIPARC:UPI000005786B; GB:AE002294; GB:AE002160; N1
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0267
C:Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G
Alignment Scores:
Pred. No.: 2,78e-46 Length: 867
Score: 789.00 Matches: 268
Percent Similarity: 40.5% Conservative: 148
Best Local Similarity: 26.1% Mismatches: 411
Query Match: 10.0% Indels: 200
DB: 31
US-10-701-844-1 (1-4435) x F81721 (1-867)
QY 433 TGTGCTCTTTAAATGGGGGATATGCGAGAAATCATGTTCTCTCAAGAAATTTC 492
Db 9 CysCys-----LeuCysAlaThrIleLeuSerProThrAlaIleLeu 22
QY 493 GATGGGAGACGTTAACTGTATCATTTCCCTATATAGGAGATCGATGGGACT 552
Db 23 PheGlyGlnAsp---AlaLeuAspLysSerAlaLeuIleThrLysAsnProAsnSerIle 41
QY 553 ACTGTTTTTCTCGAGAGAGTTAACATTAAATAATCTTGACAAATCTATTGCGACTTTG 612
Db 42 ValCysThrPheLeuGluAspCysThrMetGluAsnPheSerProAlaLeuSerHis 61
QY 613 -----CCTTTAAGTCTTTTGGCACTTATTAGGAGTTTACTGTTT 657
Db 62 AlaArgGlnAspAspProLeuTyrIleIleGlyAsnThr----- 74
QY 658 GGGAGAGGACACCTCGTTGACTTTTCGAGACATACATGCGGACTTCTACAAATGGGCGACTCTA 717
Db 75 -----HisAsnTrpPheValSerAsnLeuHisProSerThrAsnGluGluArgPhe 91
QY 718 AGTAATAGCGCTCTGATGAGCTGTTTACTATTAGAGGGTTTAAAGAAATTTATCTTTCC 777
Db 92 LeuLysGluLysGlyAsp-----LeuSerIleGlnAspPheArgPheLeuSerPheThr 109
QY 778 AATTGCAATTCATTACTTCCGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
Db 110 AspCysSerSer----- 113
QY 838 CCGACGACACATCTACACCGTCTAATGGTACTATTATTCTTAAACA---GATCTTTTG 894
```

Db 114 ---SerThrGluAspSerProSer-----lleLeuTyrHisLysAsnGlyGlnLeuPhe 130
Qy 895 TTACTCAATAATGAGAAGTCTCATTTATAGTAATTTAGTCTCTGGAGATGGGGAGCT 954
Db 131 LeuArgAsnAsnGlyAsnMetSerPheTyrArgAsnHisSerGluGlySerGlyGlyAla 150
Qy 955 ATAGATGCTAGAGCTTAAAGCTTCAAGGATTAAGCAAGCTTTGTGCTTCAAGAAAT 1014
Db 151 SerSerThrAspAlaLeuPheLeuGlnHisAsnTyrLeuPheThrAsnPheGluGluAsn 170
Qy 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACCAGTTTCTCTGCTATGCGTAAC 1074
Db 171 SerSerAlaLysAsnGlyGlyAlaIleGlnAlaGlnThr 183
Qy 1075 GAGGCTCTATTGCTTTGTACGAATGTTGCAGGAGTAAGAGGGGAGGATGCTGCT 1134
Db 183 ----- 183
Qy 1135 GTTCAGGATGGCAGCAGGAGTGTATCATCTACTTCAACAGAAGATCCAGTAGTAAGT 1194
Db 184 -----LeuSer 185
Qy 1195 TTTTCCAGAAAT---ACTGCGGTAGAGTTGTGAGGAAAGCTAGCCGAGTAGGAGGG 1251
Db 186 LeuSerArgAsnValSerSerLeuSerPheSerArgAsnArgAlaAsnLeuAsnGlyGly 205
Qy 1252 ATTTACTCTACGGAACTGCTTTCTGTAATATGGAATACTGTTTCTCAACAAT 1311
Db 206 AlaIleCysCysGlnAsnLeuIleCysSerGlyAsnValAsnProLeuPhePhe 223
Qy 1312 GTTGCTTCTCTGTTTACATTCCTGCTAGCAACCAACGAAGTGGACAGGCTTCTAATAG 1371
Db 224 -----Thr 224
Qy 1372 AGTAATAATTACGGAGATGGAGAGCTATCTTCTGAAGAATGGTCCGACAGCATCC 1431
Db 225 AsnAsnSerAlaLeuAsnGlyAlaIleCysCysIleAsnGlnGlnAsnLeuSerGlu 244
Qy 1432 AATAACTCTGGAATCTGCTTTCTGATGGAGGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491
Db 245 LysGlyCysLeuSerLeuAlaTyrAsnGlnGlu-----ThrLeuPheSerGlyAsnSer 262
Qy 1492 GCTGCTGGGAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTGTGCTAACTGTGGCCT 1551
Db 263 AlaLysGluLysGlyGlyAlaIleTyrThrLysHisMetValLeuArgHisAsnGlyPro 282
Qy 1552 GTACAATTTTAAGGAATATCCTAATGATGCTGGAGCGATTTATTAGGAGAACTGGA 1611
Db 283 ValSerPheValAsnAsnSerAlaLysLeuGlyGlyAlaIleAlaIleGlnSerGlyGly 302
Qy 1612 GAGCTCAGTTTATCTGCTGATTATGAGATATATTTTCGATGGGAATCTTAAAGAACA 1671
Db 303 SerLeuSerIleIleAlaGlyGlySerValLeuPheGlnAsnAsnSerCysHisPhe 322
Qy 1672 GCCAAGAGAATGCTGCCGATTTAATGGCGTAACGTGCTCTCAAGCCATTCGATG 1731
Db 323 SerAspGlnGlyThrVal-----ArgAsnAlaIleTyrLeu 334
Qy 1732 GATCGGGGGGAAATAACGACATTAAGCATTAAGCATTAAGCAGGCGCATCAGATTTCTTAAT 1791
Db 335 GluLysAsnAlaLeuLeuSerSerLeuGluAlaArgHisGly---AspIleLeuPhePhe 353
Qy 1792 GATCCCATCGAGATGGCAACGAATAACACGACGACGACGCTCTTCCAAACTTCTAAAA 1851
Db 354 AspProIleValGlnGluValSerProGluPheSerThrSerAlaLeuThrPro 373
Qy 1852 ATTAACGATGGTGAAGGATACACAGGGGATATGTTTTT----- 1890
Db 374 LeuArgIleGlnThrAsnThrAsnArgAlaValIlePhePheSerSerGluAsnLeuSerLys 393
Qy 1891 -----GCTAATGGAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAA 1935
Db 394 GluGluLysThrGluAlaAsnLeuIleSerLysIleGlnGlnProIleGluLeuGlnSer 413

Qy 1936 GGAAGGATTTGTTCTTCGTGAAAAGGCAAAATTATATCACTGAATTTCTTAAGTCACAGAGT 1995
Db 414 GlyCysLeuValLeuLysAspArgValIleLeuSerAlaProSerLeuSerGlnAlaPro 433
Qy 1996 GGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACAACCA 2052
Db 434 GlnAlaLeuLeuValMetAspValGlyThrSer----- 444
Qy 2053 CAACAGCCTCTCGCGCTTAATCACTGATCACTGATCACTGCTTCCAAATCTGCAATTTGCTTTCT 2112
Db 445 -----LeuThrThrSerSerAspLeuLysLeuThrThrLeu 456
Qy 2113 TCTTTGTTAGCAACAATGCGAGTTACGAATCTCTCAATCTCCAGCGCAAGATTCT 2172
Db 457 SerIleProLeuHisSerIle----- 463
Qy 2173 CATCTCGAGTCATTGTTAGCACAACTGCTGTTCTGTGTACAAATTAGTGGGCCT----- 2226
Db 464 -----AspThrGluAsnSerValSerIleGlnSerProThrLeu 476
Qy 2227 -----ATCTTTTGTAGGATTTGGAT---GATACAGCTTTATGATAGGTATGAT 2271
Db 477 SerIleGlnLysIlePheLeuSerAsnSerGluHisGluAsnPheTyrGluAsnValGlu 496
Qy 2272 TGGCTAGTCTTAATCAAAA---ATCAATGCTCTGAAAATTACAGTTAGGAGCTAAAGCCC 2328
Db 497 LeuLeuSerLysAspGlnLysAspIleProLeuLeuSerLeuProLysGlyLeuProHis 516
Qy 2329 CCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCCCTAAGTAGTGGCTATCAA 2388
Db 517 Pro-----AspLeuProAspGlyAsnLeuSerSerHisPheGlyTyrGln 531
Qy 2389 GGAAGCTGGAAGCTTGGCTGGGATCTTAATACAGCAATAATGGTCTTATCTCTGAAA 2448
Db 532 GlyAspTrpAsnPheSerTrpGlnThrSerAspGlnArgGlu-----ThrLeuVal 548
Qy 2449 GCTACATGAGCTAAACTGGGTATACTCTGGGCTGAGCGAGTAGTCTTTTGGTTCCA 2508
Db 549 AlaAsnTrpThrAlaAsnSerTyrIleProHisProGluArgGlnSerAlaLeuValAla 568
Qy 2509 AATAGTTTATCGGATCCATTTTATAGATATACGATCTGGCATTCAGCAATTCAGCAAGT 2568
Db 569 AsnThrLeuTrpAsnThrTyrSerAspMetGlnAlaValGlnSerMetIleAsnThrThr 588
Qy 2569 GTGGATGGCGCTCTTATTGTGAGGATTATGGGTTCTCGAGTTTCGAAATTTCTTCTAT 2628
Db 589 AlaGlnGlyGlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSerAsnLeuPheTyr 608
Qy 2629 ---CATGACCCGATGCTTTTAGGT---CAGGATATCGGTATATATAGTGGGGTTATTCC 2682
Db 609 SerHisGlyAsnSerGlyLysSerThrAspAsnTrpLysHisArgSerLeuGlyTyrLeu 628
Qy 2683 TTAGGACAAACTCCTAC---TTTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTA 2739
Db 629 PheGlyIleSerThrHisSerLeuAspHisSerPheCysLeuAlaIleGlyGlnLeu 648
Qy 2740 TTTGTTAGATCTAAGAGATTATGATGTGTCTTCCAAATCATCATCTTGATAGGATCC 2799
Db 649 PheGlyLysSerSerAspSerPheValThrSerAlaAspThrThrSerTyrIleAlaAla 668
Qy 2800 GTTTATCTATCTACCAACAAGCTTTATGTGGATCCTATTGTTGCGGAGATCGCGTTATC 2859
Db 669 Ile-----GlnThrGlnIleAlaThrSerLeuIleLysIleSerAla----- 682
Qy 2860 CGTGCTACTAGCGGTTGGGAATCAGCATATGAAACCTCATAT---ACATTCGACAG 2916
Db 683 GlnAlaCysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSerPheSerLys 702
Qy 2917 GAGCGCATGTTGTTGGGATAATAACTGTCTGCTGGAGGAGATTGGAGCGGATTACCG 2976
Db 703 GluGlyPheGlyAlaTrpHisSerValAlaValSerGlyGluIleGlyAlaSerIlePro 722


```
Db 480 GlyLeuTTPThrIleAsnTTPSerThrAspAspLys-----Lys 492
Qy 2449 GCTACA-----TGACCTAAACTGGGTATATCTCGGCCCTGAGCGAGTAGCTTCTTTG 2502
Db 493 AlathrValSerTTPAlaLysGlnSerPheAsnProThrAlaGluGlnGluAlaProLeu 512
Qy 2503 GTTCCAAATAGTTATGGGATCCATTTTAGATATACCATCTGCGCATTCAGCAATTCAA 2562
Db 513 ValProAsnLeuLeuTTPGlySerPheIleAspValArgSerPheGlnAsnPheIleGlu 532
Qy 2563 GCAAGTGTGATGGCGCTCTATTGTGCGAGGATTATGGGTTTCTGGAGTTTTCGAATTC 2622
Db 533 LeuGlyThrGluGlyAlaProTyrGluLysArgPheTTPValAlaGlyIleSerAsnVal 552
Qy 2623 TTCTATCATGACCGGATGCTTTAGTCAGGATATCGGTATATATAGTGGGGTTATTCC 2682
Db 553 LeuHisArgSerGlyArgGluAsnGlnArgLysPheArgHisValSerGlyGlyAlaVal 572
Qy 2683 TTAGGAGCAAACTCTACTTT--GGATCATCGATGTTTGGTCTAGCATTTACCGAAGTA 2739
Db 573 ValGlyAlaSerThrArgMetProGlyGlyAspThrLeuSerLeuGlyPheAlaGlnLeu 592
Qy 2740 TTTGTAGATCTAAAGATTATGATGTGTCTCCATCATCATCTGCTTGCATAGATCC 2799
Db 593 PheAlaA-GAspLysAspTyr-PheMetAsnThrAsnPheAlaLysThrTyrAlaGlySer 612
Qy 2800 GTTTATCTATCTACCCAAACAGCTTTATGTGGATCCTATTTTTCGGAGATGCGTTTATC 2859
Db 613 LeuArgLeu----- 615
Qy 2860 CQTGCTAGTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTTCAGAGGAG 2919
Db 616 -----GlnHisAspAlaSerLeuTyrSerValVal----- 625
Qy 2920 AGCGATGTTCTGTGGGATAATAACTGTCTGGCTGAGAGATTTGGAGCGGATACCGATT 2979
Db 626 -----SerIleLeuLeuGlyGlu-----GlyGlyLeuArgGlu 636
Qy 2980 GTGATTACTCCATCTAAGCTCTATTGTAATGAGTTGCGTCCT-----TTCGTGCAA 3030
Db 637 IleLeuLeuPro-----TyrValSerLysThrLeuProCysSerPheTyrGlyGln 653
Qy 3031 GCTGAGTTTCTTATGCGCATCAATCTTTTACAGAGAA----- 3072
Db 654 LeuSerTyrGlyHisThrAspHisArgMetLysThrGluSerLeuProProArg 673
Qy 3073 -----GCCGATCAAGCTCGGCAATTCAGAGCGGACATCTC 3108
Db 674 ArgSerArgArgIleIleLeuLeuGlyGluAspMetSerGlyLeuGluSerTTPGluLeu 693
Qy 3109 ---CTAAATCTATCAGTTCTGTTGGAGTGAAGTTTGTATCGATGTTCTAGTACACAT 3162
Db 694 GluLeuLeuLeuLysIleProAlaAlaGluAspPheSerLys---SerThrLeuHis 711
RESULT 32
p27076
polymorphic outer membrane protein h family - Chlamydoiphila pneumoniae (strain CWL029)
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: G72076
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-978 <ARN>
A:Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000046ABB; GB:AB001629; GB:AB001363; NID
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp.14
C:Superfamily: Polymorphic membrane protein H family
```

```
Alignment Scores: 2.8e-40 Length: 978
Pred. No.: 702.00 Matches: 288
Score: 38.8% Conservative: 149
Percent Similarity: 25.6% Mismatches: 419
Best Local Similarity: 8.9% Indels: 269
Query Match: 2 Gaps: 44
Db: 2

US-10-701-844-1 (1-4435) x G72076 (1-978)
Qy 405 CTTCCTTCAATGATTCTAGCTTATCTTCTGCTGCTCTTTAAATGGGGGGGATATGCAGC 464
Db 3 LeuSerPheLysSerSerPheCysLeuLeu-----AlaCysLeuCysSer 18
Qy 465 AGAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTA 524
Db 19 -----AlaSerCys-AlaPheAlaG1 25
Qy 525 TACTGTTATAGGAGATCGAGTGGGACTACTGTTTTTCTGCAGAGAGTTAAACATTAA 584
Db 25 uThrArgLeuGlyGlyAsnPheValProIleThrAsnGlnGlyGluGluLeuLe 45
Qy 585 AAATCTTGAC-----AATCTATTGACAGCTTTTGCTTTTAAGTTGTTTGGAA 632
Db 45 uThrSerAspPheValCysSerAsnPheLeuGlyAlaSerPheSerSerPheIleAs 65
Qy 633 CTTATTAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTCGAGAGACATACG 692
Db 65 nSerSerSerAsnLeuSerLeuGlyLysGlyLeuSerLeuThrPheThrSerCysG1 85
Qy 693 GACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGA 752
Db 85 nAlaProThrAsnSerAsnTyrAlaLeuLeuSerAlaAlaGluThrLeu---ThrPheL 104
Qy 753 GGGTTTTAAAGAAATTATCTCTTTTCCAAATTCAATTCTTCTGCGTACTGCTGCTGCTG 812
Db 104 AsnPheSerSerIleAsnPheThr----- 112
Qy 813 AACGACTAATAAGGATAGCCAGACTCCGACGACAAACATCTACCGCTTAATGACTACTAT 872
Db 113 -----GlyAsnGln-----SerThrGlyLeuGlyGlyLeuI1 123
Qy 873 TTATCTTAAACACATCTTTGTTTACTCAATAATAGAGAGTTCTCATTCTATAGTAATTT 932
Db 123 eTyrGlyLys---AspIleValPheGlnSerIleLysAspLeuIlePheThrThrAsnAr 142
Qy 933 AGTCTCTGGAGATGGGCGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAA 992
Db 142 gVal-----AlaTyrSerProAlaSerValThrThrSerAlaThrProAl 157
Qy 993 GCTTTGTGTCTTCCAGAAAATATCTGCTCAAGCTGATGGGGGAGCTGTCAAGTAGTCAAC 1052
Db 157 aile-----ThrThrValThrThrGlyAlaSerAlaLeuGlnProThrAs 172
Qy 1053 CAGTTTCTCTGCTATGCTTAACGAGGCTCTATTGCTTTTCTAGGAGATGTTCAGGAGT 1112
Db 172 pSerLeuThrValGluAsnIleSerGlnSerIleLysPhePheGlyAsnLeuAlaAsnPh 192
Qy 1113 AAGAGGGGGGAGGATGCTGCTGTTTCAGATGGGCGAGGAGGTCTCATCTACTTCTTC 1172
Db 192 eGly-----SerAlaIleSerSe 198
Qy 1173 AACAGAGATCCAGTAGTAGTTTTCAGAAAATATCTGCG---GTAGAGTTTGTATGGGAA 1229
Db 198 rSerProThrAlaValValLysPheIleAsnAsnThrAlaThrMetSerPheSerHisAs 218
Qy 1230 CGTAGCCGAGTAGAGGAGGG---ATTACTCTACGGGAGACGTTGTTCTCTCGAATAA 1286
Db 218 nPheThrSerSerGlyGlyGlyValIleTyrGlySerSerLeuLeuPheGluAsnAs 238
Qy 1287 TGGAAAAACCTGTTTCTCAACAATGTTGCTTCTCTCTCTTACATTTGCTGTGCTGCTG 1346
```

Db 238 nSerGlyCysIlellePheThrAlaAasnSerCysValaasnSerLeuLysGlyValThrPr 258
 QY 1347 AACAGTGCAGAGCTTCTTAATACGAGTAATAATACGAGATGGAGGCTATCTTCG 1406
 Db 258 oSerSerGlyThrTyAla-----LeuGlySerGlyAlaIleCysIl 273
 QY 1407 TAAGAAATGTCGCAAGCAGGATCCAAATCTCTGATCAGTTCTCTTGAT----- 1458
 Db 273 eProThrGlyThrPheGluLeuLysAasnGlnGlyLysCysThrPheSerTyAasnGl 293
 QY 1459 -----GGAGAGG 1466
 Db 293 yThrProAasnApsAlaGlyAlaIleTyAlaGluThrCysAasnIleValGlyAasnGlnGl 313
 QY 1467 AGTAGTTTCTTAGTAGCAATAGCTCTGCGGAAGGGGAGCTATTATGCCAAAA 1526
 Db 313 yAlaLeuLeuLeuApsSerAasnThrAlaAlaArgAasnGlyAlaIleCysAlaLysVa 333
 QY 1527 GCTCTCGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGCAATATCGCTAATGATGCTG 1586
 Db 333 lLeuAasnIleGlnGlyArgGlyProIleGluPheSerArgAasnArgAlaGluLysGlyGl 353
 QY 1587 AGCGATTATTATAGGA-----GAATCTGGAGAGCTCAGTTT 1622
 Db 353 yAlaIlePheIleGlyProSerValGlyAaspProAlaLysGlnThrSerThrLeuThrIl 373
 QY 1623 ATCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACACGCCAAAGAA 1682
 Db 373 eLeuAlaSerGluGlyAaspIleAlaPheGlnGlyAasnMetLeuAasnThrLysPro----- 391
 QY 1683 TGCTGCCGATGTTAATGGCTTAAGTGTCTCACAAGCCATTTTCGATGGATCGGAGG 1742
 Db 392 -----GlyIle-----ArgAasnAlaIleThrValGluAlaGlyGl 403
 QY 1743 GAAATAACGACATTAAAGACTAAAGCAGGCATCAGATTCTCTTTAATGATCCCATCGA 1802
 Db 403 yGluIleValSerLeuSerAlaGlnGlyGlySerArgLeuValPheTyAaspProIleTh 423
 QY 1803 GATGGCAACGGAATAACACGCGCAGTCTCCAACTCTTAACTTCTTAAATAACGATGG 1862
 Db 423 rHisSerLeuProThrThrSerPro-----SerAasnLysAaspIleThrIleAasn---Al 440
 QY 1863 TGAAGNATCACAGGGGATATCTTTT----- 1890
 Db 440 aAasnGlyAlaSerGlySerValValPheThrSerLysGlyLeuSerSerThrGluLeuLe 460
 QY 1891 -----GCTAATGGAAGCAGTACTTTGTATACCAAAATGTTACGATAGCAAGGAAGAT 1943
 Db 460 uLeuProAlaAasnThrThrThrIleLeuLeuGlyThrValLysIleAlaSerGlyGluLe 480
 QY 1944 TGTCTTTCGTGAAAGCAAAATATCAGTGAATCTCTAAGTCAGACAGGT---GGGAG 2000
 Db 480 uLysIleThrAaspAasnAlaValValAasnValLeuGlyPheAlaThrGlnGlySerGlyGl 500
 QY 2001 TCTGTATATGCAAGCTGGAGTACATGGGATTTGTAACTCCCAACACCAACACAGCC 2060
 Db 500 nLeuThrLeuGlySerGlyGlyThrLeuGlyLeuAlaThr-----ProThrGlyAl 517
 QY 2061 TCTGCGCGCTAATCAGTTGTATGATCGCTTTCCAACTCTGCAATTTGTCTTCTTCTTTGTT 2120
 Db 517 aProAlaAlaValAasp---PheThrIleGlyLysLeuAlaPheAaspProPheSerPheLe 536
 QY 2121 AGCAAAACAAATGACAGTACGAATCTCTTACCAATCTCCAGCGCAAGATCTCATCTCGC 2180
 Db 536 uLysArgAaspPheValSer----- 542
 QY 2181 AGTCATTTGGTAGCACTGCTGGT-----TCTGTTACAATTAGTGGGCTATCTTTT 2234
 Db 543 -----AlaSerValAasnAlaGlyThrLysAasnValThrLeuThrGlyAlaLeuValLe 560
 QY 2235 TGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2294
 Db 560 uAaspGluHisAaspValThr-----AaspLeuTyAasp----- 570

QY 2295 CAATGCTCTGAAATTTACAGTTTAGGACTAAGCCCCAGCTAATGCCCCATCAGATTTCAC 2354
 Db 571 -----MetValSerLeuGlnSerProValAlaIleProIleAlaValPh 585
 QY 2355 TCTAGGGAATGAGATGCTTAAG----- 2376
 Db 585 eLysGlyAlaThrValThrLysThrGlyPheProAaspGlyGluIleAlaThrProSerHi 605
 QY 2377 -TATGGCTATCAAGCAAGCTCGAAGCTTGGTGGGAT----- 2412
 Db 605 sTyGlyTyGlnGlyLysTrpSerTyThrTrpSerArgProLeuLeuLeuProAlaPr 625
 QY 2413 -----CCTAATACAGCAATAATGCTCTTATCTCTGAAGCTACATGG-- 2457
 Db 625 oAaspGlyGlyPheProGlyGlyProSerProSerAlaAasnThrLeuTyAlaValTrpAs 645
 QY 2458 -----ACTAAACCTGGGTATATCTCTGGGCTGAGCAGTAGCTTCTTGGT 2504
 Db 645 nSerAaspThrLeuValArgSerThrTyIleLeuAaspProGluArgTyGlyGluIleVa 665
 QY 2505 TCCAATAGTTTATGGGATCCATTTTAGATATAGATATAGCTGCGCATTCAGCAATTCAA-- 2562
 Db 665 lSerAasnSerLeuTrpIleSerPheLeuGlyAasnGlnAlaPheSerAaspIleLeuGlnAs 685
 QY 2563 ----GCAAGTGTGGATGGGCGCTCTTATGTCGAGGATTTATGGTTCGGAGTTTCGAA 2618
 Db 685 pValLeuLeuIleAaspHis-----ProGlyLeuSerIleThrAla----- 698
 QY 2619 TTTCTTCTATCAGACCGGATGCTTTAGT-----CAGGG 2654
 Db 699 -----LysAlaLeuGlyAlaTyValGluHisThrProArgGlnGl 712
 QY 2655 ATATCGGTATATAGTGGGGTTAT-----TCTTAGGAGCAAACTCCTA 2699
 Db 712 yHisGluGlyPheSerGlyArgTyGlyGlyTyGlnAlaAlaLeuSerMetAasn---Ty 731
 QY 2700 CTTTGGATCATCGATGTTTGGTCTAGCATTTCAGCAAGTATTGTTAGTAGTCTTAAAGATTA 2759
 Db 731 rThrAaspHisThrThrLeuGlyLeuSerPheGlyGlnLeuTyGlyLysThrAasn----- 749
 QY 2760 TGTAGTGTCTGTTCCAAATCATCATGTTGATGATGATGATGATGATGATGATGATGATGAT 2819
 Db 750 -----AlaAasnProTyAaspSerArgCysSerGluGlnMetTyLeuLeuSer----- 765
 QY 2820 AGCTTATGTCGATCTTATTTCTCGA-----CATGC 2852
 Db 766 -----PhePheGlyGlnPheProIleValThrGlnLysSerGluAl 779
 QY 2853 GTTTATC-----CGTGTAGCTACGGGTTTGGGAATCAGCATATGAAAACTCATAT-- 2904
 Db 779 aLeuIleSerTrpLysAlaAlaTyGlyTySerLysAasnHisLeuAasnThrThrTyLe 799
 QY 2905 -----ACATTTGCAGAGGAGCGATGTTGTTGGGATAATAAATGCTGCTGCTGGAGA 2957
 Db 799 uArgProAaspLysAlaProLysSerGlnGlyGlnTrpHisAasnAasnSerTyTyValle 819
 QY 2958 GATTGAGCGGATTAACCGAT-----GTGATTATCTCATTAAGCTCTATT 3005
 Db 819 uIleSerAlaGluHisProPheLeuAasnTrpCysLeuLeuThrArgProLeuAlaGlnAl 839
 QY 3006 GAATGATGTCGCTCTTCTGTCGCAAGCTGAGTTTCTTATGCGGATCATGAACTCTTTTAC 3065
 Db 839 aTrpAaspLeuSerGlyPheIleSerAlaGluPheLeuGlyGlyTrpGlnSerLysPheTh 859
 QY 3066 AGAGGAAGCGGATCAAGCTCGGGCATTCAGAGCGGACATCTCTTAATATCATGATTC 3125
 Db 859 xGluThrGlyAaspLeuGlnArgSerPheSerArgGlyLysGlyTyAasnValSerLeuPr 879
 QY 3126 TGTGGAGTGAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3164
 Db 879 oIleGly-----CysSerSerGlnTrpPheThrProPheLysLysAlaPr 894

```
QY 3165 TAATAAATAGCTTTATGGCGCTTATATCTGTGATCTTATCGCACCATCTCTGGTAC 3224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 894 oSerThrLeuThrIleLeuLeuAlaValThrProAspIleValArgValAsnProHisAs 914
QY 3225 TGAGACAACGCTCTATCCATCAAGAGACATGGACAACAGATGCTCTTCATTTAGCAAG 3284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 nIleValThrValSerAsnGlnGluSerThrSerIleSerGlyAlaAsnLeuArgAr 934
QY 3285 ACATCGAGTGTGGTTAGAGATCTATGTATGCTCTCTCAACAGTAATATAGAAATATA 3344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 934 gHisGlyLeuPheValGlnHisAspValValAspLeuThrGluAspThrGlnAlaPh 954
QY 3345 TGGCCATGGAAGATATGATATCGAGATGCTCTCGAGGCTATGTTGAGTGCAGGAAG 3404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 954 eLeuAsnTyrThrPheAspGlyLeuAsnGlyPheThrAsnHisArgValSerThrGlyLe 974
QY 3405 TAGAGTCCGGTTC 3417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 974 uLysSerThrPhe 978

RESULT 33
C86547
polymorphic outer membrane protein H family [imported] - Chlamydothila pneumoniae (strain
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86547
R:Shirai, M.; Hiraoka, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: C86547
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-978 <STO>
A:Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI000131C83; GB:BA000008; NID:g8978825; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_14
C:Superfamily: Polymorphic membrane protein H family

Alignment Scores:
Pred. No.: 6.19e-40 Length: 978
Score: 697.00 Matches: 287
Percent Similarity: 38.8% Conservative: 150
Best Local Similarity: 25.5% Mismatches: 419
Query Match: 8.8% Indels: 269
DB: 2 Gaps: 44

US-10-701-844-1 (1-4435) x C86547 (1-978)
QY 405 CTTTCTTCAATGATCTAGCTTATCTTGCTCTTTTAAATGGGGGGGATATGCAGC 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LeuSerPheLeuSerSerSerPheCysLeuLeu-----AlaCysLeuCysSer 18
QY 465 AGAAATCATGGTCTCTCAAGGAATTTACGATGGGAGAGCTTAACGTATCATTTCCCTA 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 19 -----AlaSerCys-AlaPheAlaGl 25
QY 525 TACTGTATPAGGATCCGAGTGGGACTACTGTTTTCTCGCAGGAGAGTTAAACATTAA 584
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 uThrArgLeuGlyGlyAsnPheValProProlleThrAsnGlnGlyGluGluIleLeu 45
QY 585 AAATCTTGAC-----AATTCTATTCCAGCTTTGCCCTTTAAATGTTGTTGGAA 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 uThrSerAspPheValCysSerAsnPheLeuGlyAlaSerPheSerSerPheIleAs 65
QY 633 CTTATATGGAGTTTACTGTTTATGGAGAGACACTGTTTCATCTTCGAAACATACG 692
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 nSerSerSerAsnLeuSerLeuGlyGlySerLeuThrPheThrSerCysGl 85
QY 693 GACTTCTACAAATGGGGCAGCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTGA 752
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 nAlaProThrAsnSerAsnTyrAlaLeuLeuSerAlaAlaGluThrLeu---ThrPheLy 104
```

```
QY 753 GGGTTTTAAAGAATTATCCTTTTTCCAATTGCAATTCAATTACTTGCCTGCTGCTGCTGC 812
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 104 sAsnPheSerSerIleAsnPheThr----- 112
QY 813 AACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACCGTCTAATGGTACTAT 872
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 -----GlyAsnGln-----SerThrGlyLeuGlyGlyLeuI 123
QY 873 TTATTTCTAAACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTTCTATAGTAATTT 932
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 eTyrGlyLys---AspIleValPheGlnSerIleLysAspLeuIlePheThrThrAsnAr 142
QY 933 AGTCTCTGGAGATGGGGAGGCTATAGATGCTAAAGACTTAACGGTTCAAGGAATTAGCA 992
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 gVal-----AlaTyrSerProAlaSerValThrThrSerAlaThrProAl 157
QY 993 GCTTTGTCTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCAC 1052
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 157 aIle-----ThrThrValThrThrGlyAlaSerAlaLeuGlnProThrAs 172
QY 1053 CAGTTTCTCTGCTATGCTAACGAGGCTCTTATTTCTAGCGAATGTTGCGAGGAGT 1112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 pSerLeuThrValGluAsnIleSerGlnSerIleLysPhePheGlyAsnLeuAlaAsnPh 192
QY 1113 AAGAGGGGAGGAGTGTCTGCTTTCAGGATGGGAGGAGGAGTGTCTCATCTACTTCTC 1172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 eGly----- 198
QY 1173 AACAGAGATCCAGTGTAGTATTTTCCAGAAATACTGCG---GTAGAGTTTGTAGGAA 1229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 198 rSerProThrAlaValValLysPheIleAsnAsnThrAlaThrMetSerPheSerHisAs 218
QY 1230 CGTAGCCCGAGTAGGAGGAGG---ATTACTCTTCAGCGAACGTTGTTCTCTGGAATAA 1286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 nPheThrSerSerGlyGlyGlyValIleTyrGlySerSerLeuLeuPheGluAsnAs 238
QY 1287 TGGAAAAACCTGTTTCTCAACAATGTTGCTTCTCCTGTTTACATGCTGCTAGCAACC 1346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 nSerGlyCysIleIlePheThrAlaAsnSerCysValAsnSerLeuLysGlyValThrPr 258
QY 1347 AACAGAGTGCAGAGGCTCTAATACGAGTAATAATTACGAGATGAGGAGGCTATCTTCG 1406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 258 oSerSerGlyThrTyrAla-----LeuGlySerGlyGlyAlaIleCysI 273
QY 1407 TAAGAATGTGCGCAAGCAGGATCCAATAACTCTGGATCAGTTTCTCTTGAT----- 1458
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 eProThrGlyThrPheGluLeuLysAsnAsnGlnGlyLysCysThrPheSerTyrAsnGl 293
QY 1459 -----GGAGAGGG 1466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 yThrProAsnAspAlaGlyAlaIleTyrAlaGluThrCysAsnIleValGlyAsnGlnGl 313
QY 1467 AGTAGTTTCTTAGTAGCAATGTAGCTGCGGAAAGGGGAGCTATTTATGCCAAAAA 1526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 yAlaLeuLeuLeuAspSerAsnThrAlaAlaArgAsnGlyGlyAlaIleCysAlaLysVa 333
QY 1527 GCTCTCGGTTGCTAACTGTGCGCTGTACAAATTTTAAAGGAATATCGCTAATGATGGTG 1586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 333 lLeuAsnIleGlnGlyArgGlyProIleGluPheSerArgAsnArgAlaGluLysGlyGl 353
QY 1587 AGCGATTATTATAGA-----GAATCTGGAGAGCTCAGTTT 1622
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 yAlaIlePheIleGlyProSerValGlyAspProAlaLysGlnThrSerThrLeuThrIl 373
QY 1623 ATCTGCTGATTATGAGATATTTTTCGATGGGAATCTTAAAGAAAGACAGCAAGAGAA 1682
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 373 eLeuAlaSerGluGlyAsnIleAlaPheGlnGlyAsnMetLeuAsnThrLysPro----- 391
QY 1683 TGCTGCCGATGTTAATGGCGTAACGTGTCTCTCAAGCCCATTTTCGATGGGATCGGGAG 1742
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 392 -----GlyIle-----ArgAsnAlaIleThrValGluAlaGlyGl 403
```


1743 GAAATAACACGACATTAAGAGCTAAAGCAGCGCATCAGATTCTCTTAAATGATCCCATCGA 1802
 403 yGluIleValSerLeuSerAlaGlnGlySerArgLeuValPheTyrAspProIleTh 423
 1803 GATGGCAACCGAAATAACACGAGCGAGCTTCTCAAACTTCTAAAAATTAAACGATGG 1862
 423 rHisSerLeuProThrThrSerPro-----SerAsnLysAspIleThrIleAsn--Al 440
 1863 TGAAGATACACAGGGGATATGTTT 1890
 440 aAsnGlyAlaSerGlySerValValPheThrSerLysGlyLeuSerSerThrGluLeu 460
 1891 -----GCTAAAGGAGCAGTAGTCTTGTACCAAAATGTTACCATAGAGCAAGAGGAT 1943
 460 uLeuProAlaAsnThrThrThrIleLeuLeuGlyThrValLysIleAlaSerGlyGlu 480
 1944 TGTCTTCTCGTGAAGGCAAAATATCATGAGTAATCTCTTAAGTCAGACAGGT---GGGAG 2000
 480 uLysIleThrAspAsnAlaValValAsnValLeuGlyPheAlaThrGlnGlySerGlyG 500
 2001 TCTGTATATGAAGCTGGGAGTACATGGGATTTGTAACTCCCAACACCACCAACAGCC 2060
 500 nLeuThrLeuGlySerGlyThrLeuGlyLeuAlaThr-----ProThrGlyAl 517
 2061 TCCTGCCGCTAATCAGTTGATCAGCTTCCCAATCTGCAATTTGCTCTTCTTCTTGT 2120
 517 aProAlaAlaValAsp---PheThrIleGlyLysLeuAlaPheAspProPheSerPhe 536
 2121 AGCAAAACAATCAGTACGAACTCTCTACCAATCTCTCCAGCGCAAGATTCTCATCTGC 2180
 536 uLysArgAspPheValSer----- 542
 2181 AGTCATTGGTAGCAAACTGCTGT-----TCTGTTCAATAGTGGGCTATCTTTT 2234
 543 -----AlaSerValAsnAlaGlyThrLysAsnValThrLeuThrGlyAlaLeuVal 560
 2235 TGAGGATTTGGATGATCAGCTTATGATGATGATGATGATGATGATGATGATGATGAT 2294
 560 uAspGluHisAspValThr-----AspLeuTyrAsp----- 570
 2295 CAATGCTCTGAATAATACAGTATAGGAGTAAGCCCACTAATGCCCATCAGATTTGAC 2354
 571 -----MetValSerLeuGlnSerProValAlaIleProIleAlaValP 585
 2355 TCTAGGGAATGAGATGCTAAG----- 2376
 585 eLysGlyAlaThrValThrLysThrGlyPheProAspGlyGluIleAlaThrProSerHi 605
 2377 -TATGGCTATCAAGGAAGCTGGAAGCTTGGTGGAT----- 2412
 605 sTyrGlyTyrGlnGlyLysTrpSerTyrThrTrpSerArgProLeuLeuIleProAlaPr 625
 2413 -----CCTAATACAGCAATAATGGTCTTATCTCTGAAGAGCTACATGG-- 2457
 625 oAspGlyPheProGlyGlyProSerProSerAlaAsnThrLeuTyrAlaValTrpAs 645
 2458 -----ACTAAACTGGGTATAATCTCTGGGCTCAGCGAGTAGCTTCTTGGT 2504
 645 nSerAspThrLeuValArgSerThrTyrIleLeuAspProGluArgTyrGlyGluIleVa 665
 2505 TCCAAATAGTTATGGGATCCATTTAGATATACGATCTCGGCATTCAGCAATTCAA-- 2562
 665 lSerAsnSerLeuTrpIleSerPheLeuGlyAsnGlnAlaPheSerAspIleLeuGlnAs 685
 2563 ----GCAAGTGGAGGGCGCTCTTATTTGTCGAGGATTATGGTTCCTGGAGTTTCGA 2618
 685 pValLeuLeuIleAspHis-----ProGlyLeuSerIleThrAla----- 698
 2619 TTTCTTCTATCATGACCGCGATGCTTTAGT-----CAGGG 2654
 699 -----LysAlaLeuGlyAlaTyrValGluHisThrProArgGlnGl 712
 2655 ATATCGGTATATTAGTGGGGTTAT-----TCCTTAGGAGCAAACTCTCTA 2699

712 yHisGluGlyPheSerGlyArgTyrGlyGlyTyrGlnAlaLeuSerMetAsn---Ty 731
 2700 CTTTGGATCATCGATTTTGGTGTAGCATTTTACCGAAGTATTGTTGTAGATCTTAAAGATTA 2759
 731 rThrAspHisThrThrLeuGlyLeuSerPheGlyGlnLeuTyrGlyLysThrAsn----- 749
 2760 TGTAGTGTGTCGTTCCCAATCATCATGCTTGCAATAGGATCCGTTTATCTATCTACCAACA 2819
 750 -----AlaAsnProTyrAspSerArgCysSerGluGlnMetTyrLeuLeuSer----- 765
 2820 AGCTTTATGTGATCTATTGTTCCGA-----GATGC 2852
 766 -----PhePheGlyGlnPheProIleValThrGlnLysSerGluAl 779
 2853 GTTTATC-----CGTGTAGCTACGGGTTTGGGAATCAGCATATGAAAACTCATAT-- 2904
 779 aLeuIleSerTrpLysAlaAlaTyrGlyTyrSerLysAsnHisLeuAsnThrThrTyrLe 799
 2905 -----ACATTTGCAGAGGAGGAGCGATGCTGTTGGGATAAATCACTGCTGCTGGGAGA 2957
 799 uArgProAspLysAlaProLysSerGlnGlyGlnTrpHisAsnAsnSerTyrTyrValle 819
 2958 GATTGAGCGGATTAACGATT-----GTGATTACTCCATCTAAGCTCTATTT 3005
 819 uIleSerAlaGluHisProPheLeuAsnTrpCysLeuLeuThrArgProLeuAlaGlnAl 839
 3006 GAATGATGTCGCTCTTCTGTCGAAGCTGAGTTTCTTATCGCGATCATGAATCTTTTAC 3065
 839 aTrpAspLeuSerGlyPheIleSerAlaGluPheLeuGlyGlyTrpGlnSerLysPheTh 859
 3066 AGAGGAAGCGCATCAAGCTCGGGCATTCAAGAGCGGACATCTCTAAATCTATCAGTTCC 3125
 859 rGluThrGlyAspLeuGlnArgSerPheSerArgGlyLysGlyTyrAsnValSerLeuPr 879
 3126 TGTGAGGAGAAATTTGATGATGCTTCTAGT-----ACATATCC 3164
 879 oIleGly-----CysSerSerGlnTrpPheThrProPheLysLysAlaPr 894
 3165 TAATAATATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTAC 3224
 894 oSerThrLeuThrIleLysLeuAlaTyrLysProAspIleTyrArgValAsnProHisAs 914
 3225 TGAGACAAGCTCTCTATCCATCAAGACATCGACAACAGATGCGCTTTTCAATTTAGCAAG 3284
 914 nIleValThrValValSerAsnGlnGluSerThrSerIleSerGlyAlaAsnLeuArgAr 934
 3285 ACATGAGGATTTGTTAGAGGATCTATGTATGCTTCTCTTAAACAGATTAATATAGATATA 3344
 934 gHisGlyLeuPheValGlnIleHisAspValValAspLeuThrGluAspThrGlnAlaPh 954
 3345 TGCGCATGGAAGATATGAGTATCGAGATCGATGCTTCTCGAGGCTATGTTGATGTCGGAAG 3404
 954 eLeuAsnThrPheAspGlyLysAsnGlyPheThrAsnHisArgValSerThrGlyLe 974
 3405 TAGAGTCCGGTTC 3417
 974 uLysSerThrPhe 978
 RESULT 34
 B81593
 polymorphic membrane protein H family CP0298 [imported] - Chlamydothila pneumoniae (str
 C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C;Accession: B81593
 C;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
 Nucleic Acids Res. 28, 1397-1406, 2000
 A;Title: Genome sequence of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A;Reference number: B81593
 A;Accession: B81593
 A;Status: preliminary
 A;Molecule type: DNA

A/Residues: 1-978 <REA>
A/Cross-references: UNIPROT:Q92895; UNIPARC:UPI0000131C83; GB:AE002191; GB:AE002161; NID
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: CP0298
C/Superfamily: Polymorphic membrane protein H family

```
Db      543  -----AlaSerValAsnAlaGlyThrLysAsnValThrLeuThrGlyAlaLeuValLe 560
QY      2235  TGAGATTGGATGATACAGCTTATGATAGATTGATGGCTAGGTTCTTAATCAAAAAAT 2294
Db      560  uAspGluHisAspValThr-----AspLeuTyrAsp----- 570
QY      2295  CAATGCTCCCTGAAATTACATCTAGGAGCTAGGAGCCCGCAGCTAATGCCCCATCAGATTGAC 2354
Db      571  -----MetValSerLeuGlnSerProValAlaIleProIleAlaValPh 585
QY      2355  TCTAGGGAATGAGATGCCTAAG----- 2376
Db      585  eLysGlyAlaThrValThrLysThrGlyPheProAspGlyGluIleAlaThrProSerHi 605
QY      2377  -TATGGCTATCAAGGAGCTGGAGCTCGCTGGGAT----- 2412
Db      605  sTyrGlyTyrGlnGlyLysTyrThrTyrSerArgProLeuLeuIleProAlaPr 625
QY      2413  -----CCTAATACAGCAATAATGCTCTTATCTCTGAGGCTAGCTTTCTTTGGT 2457
Db      625  oAspGlyGlyPheProGlyGlyProSerProSerAlaAsnThrLeuTyrAlaValTyrPAs 645
QY      2458  -----ACTAAACCTGGTATAATCTCTGGGCTCGAGCGAGTAGCTTTCTTTGGT 2504
Db      645  nSerAspThrLeuValArgSerThrTyrIleLeuAspProGluArgTyrGlyGluIleVa 665
QY      2505  TCCAAATAGTTTATGGGATGCCATTTTATGATATACGATCGCCATCGCAATTCAA-- 2562
Db      665  lSerAsnSerLeuTyrIleSerPheLeuGlyAsnGlnAlaPheSerAspIleLeuGlnAs 685
QY      2563  ---GCAAGTGTGGTGGCGCTCTTATTCTCGAGGATATGCTTCTGGAGTTTCGAA 2618
Db      685  pValLeuLeuIleAspHis-----ProGlyLeuSerIleThrAla----- 698
QY      2619  TTTCTTCTATCATGACCGCATGCTTTAGGT-----CAGGG 2654
Db      699  -----LysAlaLeuGlyAlaTyrValGluHisThrProArgGlnGI 712
QY      2655  ATATCGGTATATTAGTGGGGTTAT-----TCCTTAGGAGCAAACTCCTTA 2699
Db      712  yHisGluGlyPheSerGlyArgTyrGlyGlyTyrGlnAlaAlaLeuSerMetAsn---Ty 731
QY      2700  CTTTGGATCATCATGTTTGGCTAGCAATTACGAAGTATTTGCTAGATCTAAAGATTA 2759
Db      731  rThrAspHisThrThrLeuGlyLeuSerPheGlyGlnLeuTyrGlyLysThrAsn----- 749
QY      2760  TGTAGTGTGCTTCCCAATCATCATGCTTCATAGGATCGTGTATCTATCTACCCAAACA 2819
Db      750  -----AlaAsnProTyrAspSerArgCysSerGluGlnMetTyrLeuLeuSer----- 765
QY      2820  AGCTTTATGTGGATCCTATTGTTGGGA-----GATGC 2852
Db      766  -----PhePheGlyGlnPheProIleValThrGlnLysSerGluAl 779
QY      2853  GTTTATC-----CGTCTAGCTACGGTTGGGAATCAGCATATGAAACCTCATAT-- 2904
Db      779  aLeuIleSerTrpLysAlaIaTyrGlyTyrSerLysAsnHisLeuAsnThrThrTyrLe 799
QY      2905  -----ACATTTGCAGAGGAGCGCATGTTCTGTTGGGATAATAACTGTCTGGCTGGAGA 2957
Db      799  uArgProAspLysAlaProLysSerGlnGlyGlnTrpHisAsnAsnSerTyrTyrValLe 819
QY      2958  GATTGGAGCGGGATATCCGATTT-----GTGATTACTCCATCTCAAGCTCTATTT 3005
Db      819  uIleSerAlaGluHisProPheLeuAsnTyrCysLeuLeuThrArgProLeuAlaGlnAl 839
QY      3006  GAATGAGTGTGGTCTCTTCTGCGAAGCTGAGTTTCTTATGCCCATCATGAATCTTTTAC 3065
Db      839  aTrpAspLeuSerGlyPheIleSerAlaGluPheLeuGlyGlyTyrGlnSerLysPheTh 859
QY      3066  AGAGGAAGCGCATCAAGCTGGGATTCMAAGAGCGGACATCTCTCTAAATCTATCAGTTCC 3125
```

```
Db      859  rGluThrGlyAspLeuGlnArgSerPheSerArgGlyLysGlyTyrAsnValSerLeuPr 879
QY      3126  TGTGGAGTGAAGTTTGATCGATGTTCTAGT-----ACATATCC 3164
Db      879  oIleGly-----CysSerSerGlnTrpPheThrProPheLysLysAlaPr 894
QY      3165  TAATAAATATATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTAC 3224
Db      894  oSerThrLeuThrIleLysLeuAlaTyrLysProAspIleTyrArgValAsnProHisAs 914
QY      3225  TGAGACAACGCTCCATCCCATCAAGACATGGACACACATGCTTTCATTTAGCAAG 3284
Db      914  nIleValThrValValSerAsnGlnLysSerIleSerGlyAlaAsnLeuArgAr 934
QY      3285  ACATCGAGTGTGGTTAGAGGATCTATGTATGCTTCTCTAAACAAGTAATAAGAAATATA 3344
Db      934  gHisGlyLeuPheValGlnIleHisAspValValAspLeuThrGluAspThrGlnAlaPh 954
QY      3345  TGGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTGGAGTGCAGGAAG 3404
Db      954  eLeuAsnTyrThrPheAspGlyLysAsnGlyPheThrAsnHisArgValSerThrGlyLe 974
QY      3405  TAGATCTCGGTTTC 3417
Db      974  uLysSerThrPhe 978

RESULT 35
F81722
polymorphic membrane protein E/F family TC0261 [imported] - Chlamydia muridarum (strain
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81722
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Swinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: F81722
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-976 <RET>
A:Cross-references: UNIPROT:Q9PL47; UNIPARC:UPI0000057865; GB:AE002293; GB:AE002160; NI
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0261

Alignment Scores:
Pred. No.: 9,93e-39 Length: 976
Score: 679.50 Matches: 265
Percent Similarity: 40.9% Conservative: 150
Best Local Similarity: 26.1% Mismatches: 413
Query Match: 8.6% Indels: 187
DB: 2 Gaps: 38

US-10-701-844-1 (1-4435) x F81722 (1-976)
QY      613  CCTTTAAGTTGTTTGGGAACCTATTAGGAGGTTTACTGTTTAGGAGAGGACACTCG 672
Db      33  ProTyrHisMetThrGlyLeuPhePheProLysValAsnLeuLeuGlyAspThrHisAsn 52
QY      673  TTGACT-----TTCGAGAACATA-----CGGACTTCT 699
Db      53  LeuThrAspTyrHisLeuAspAsnLeuLysCysIleLeuAlaCysLeuGlnArgThrPro 72
QY      700  ACAAAATGGGGAGCT-----CTAAGTAATAGCGCTGCTGAT 735
Db      73  TyrGluGlyAlaAlaPheThrValThrAspTyrLeuGlyPheSerAspThrGlnLysAsp 92
QY      736  GGAAGTCTTTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCCAAATTCATTCATT 795
Db      93  GlyIlePheCys-----PheLysAsnLeuThr----- 101
QY      796  GCGGTACTGCTGCTGCACGACACTAATAAGGGTAGCCAGACTCCGACGACACACATCTACA 855
```

Db 102 ProGluSerGlyGlyValIleGlySerProThrGlnAsnThrProThrIle 118
 QY 856 CCGTCTAATGTACTATTATTCTAAACAGATCTTTTGTACTCAATAATAGAGAGTTC 915
 Db 119 LysIleHisAsnThrIle-----GlyProValLeuPheGluAsnAsnThrCysHis 135
 QY 916 TCATTCTATAGTAATTTA-----GTCTCGGAGATGGGGAGCT 954
 Db 136 ArgLeuTyrThrGlnThrAspProGluAsnGluGlyAsnLysAlaAargGluGlyGlyAla 155
 QY 955 ATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTGTGCTTCCAGAAAT 1014
 Db 156 IleHisAlaGlyAspValTyrIleSerAsnAsnGlnAsnLeuValGlyPheIleLysAsn 175
 QY 1015 ACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACC----- 1053
 Db 176 PheAlaTyrValGlnGlyGlyAlaIleSerAlaAsnThrPheAlaTyrLysGluAsnLys 195
 QY 1054 ---AGTTCTCTGCTATGCTTAACGAGGCTCTATTTGCTTGTAGCGAATGTTGCAGGA 1110
 Db 196 SerSerPheLeuCysLeuAsnAsnSerCys-----IleGlnThrLysThrGly 211
 QY 1111 GTAAGCGGGAGGAGTCTGCTCTCTCAGGATGGCAGGAGTCTCATCTACT 1170
 Db 212 GlyLysGlyIleValIleTyr-----ValSerThrSerCys 223
 QY 1171 TCAACAGAGATCCAGTAGTAGTATTTTCCAGAAATATCGCGTAGAGTTGATGGGAAC 1230
 Db 224 SerPheGlu-----AsnAsnAsnLysAspLeuLeuPheIleGlnAsn 237
 QY 1231 GPAGCCGAGTAGGAGGAGGATTTACTCTCTACGGGAACGTTGCTTCTGTAATATGGA 1290
 Db 238 SerGlyCysAlaGlyAlaIlePheSerProThrCysSerLeuIleGlyAsnGlnGly 257
 QY 1291 AAAACCTGTTTCTCAACATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1350
 Db 258 AspIleValPheTyrSerAsnHisGlyPheLys----- 268
 QY 1351 AGTGACAGGCTTCTAATACGAGTAGTAATAATACGAGGAGTAGGAGCTATCTCTGTAAG 1410
 Db 269 -----AsnValAspAsnAlaThrAsnGluSerGlyAspGlyGlyAlaIle-----Lys 284
 QY 1411 AATGTGCGCAAGCAGGATCCAAATACTCTGATCAGTTTCTCTTGTGAGAGGAGGATA 1470
 Db 285 ValThrThrArgLeuAspIleThrAsnAsnGlySer-----Gln 297
 QY 1471 GTTTTCTTTAGTACGATGCTGCTGGAAAGGGGAGCTATTATATGCCAAAAGCTC 1530
 Db 298 IlePhePheSerAspAsnIleSerArgAsnPheGlyGlyAlaIleHisAlaProCysLeu 317
 QY 1531 TCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAAT---GATGGTGA 1587
 Db 318 HisLeuValGlyAsnGlyProThrTyrPheThrAsnAsnIleAlaAsnHisThrGlyGly 337
 QY 1588 GCGATTTATTAGGAGAAFTCGAGAGCTCAGTTTATCTGCTGATTAAGAGATATTAT 1647
 Db 338 AlaIleTyrIleThrGlyThrGluThrSerLysIleSerAlaAspHisAlaIle 357
 QY 1648 TTCGATGGGAATCTTAAAGAACAGCAGCAAGAGATGCTGCGATGTTAATGCGCTAAT 1707
 Db 358 PheAspAsnAsnIle-----SerAlaAsnAlaThrAsnAlaAspGlySerSer 373
 QY 1708 GTGCTCTCA-----CAAGCCATTTCGATGGGATCGGAGGAGGAAATAACG 1752
 Db 374 SerAsnThrAsnProHisArgAsnAlaIleThrMetAspAsnSerAlaGlyGlyIle 393
 QY 1753 ACATTAGAGCTAAAGCAGGCGATTCAGATTCTCTTTTAATGATCCCATCGAGATGGCAAC 1812
 Db 394 GluLeuGlyAlaGlyLysSerGlnAsnLeuIlePheTyrAspProIleGlnValThrAsn 413
 QY 1813 GGAAT-----AACGACCGCGCAGCTCTTCCAACTCTTCCAACTTCTTAAATTAAC 1857

Db 414 AlaGlyValThrValAspPheAsnLysAspAlaSerGlnThrGlyCysValValPheSer 433
 QY 1858 GATCGTGAAGGATACACAGGGGATATTGTTTGTCTAATGAAGACAGTACTTTGTACCAA 1917
 Db 434 GlyAlaThrValLeuSerAlaAspIleSerGlnAlaAsnLeuGlnThrLysThrProAla 453
 QY 1918 AATGTTACGATAGCAGCAGGAGGATTTGCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 1977
 Db 454 ThrLeuThrLeuSerHisGlyLeuLeuCysIleGluAspArgAlaGlnLeuThrValAsn 473
 QY 1978 TCTCTAAGTCAGACAGGTGGGAGTCTGATATGAAAGCTGG-----AGTACATGG 2028
 Db 474 AsnPheThrGlnThrGlyIleValAlaLeuGlyAsnGlyAlaValLeuSerSerTyr 493
 QY 2029 GATTTGTTAACTCCACCAACCAACCAACAGCTCTCT-----GCCGCTAAATCAGTTG 2079
 Db 494 GlnHisSerThrThrAspAlaThrGlnThrProThrThrThrThrThrAspAlaSer 513
 QY 2080 ATCAGCTTTCCATCTGCAATTTGCTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2139
 Db 514 ValThrLeuAsnHisIleGlyLeuAsnLeuProSerIleLeuLysAspGlyAlaGluMet 533
 QY 2140 AATCCTCTCAACCAATCTCCA-----CGCAGATTTCTCATCTCGCAGTCATTTGGTAGC 2193
 Db 534 ProLeuLeuTyrValGluProIleSerThrThrGlnGlyAsnThrThrThrThrThrSer 553
 QY 2194 ACAACTGCTGTTCTGTTTACAATTTAGTGGGCTATC-----TTTTTTGAGGATTTGGAT 2247
 Db 554 AspThrAlaAlaSerPheSerLeuAsnGlyAlaThrLeuSerLeuIleAspGluAspGly 573
 QY 2248 GATCAGCTTATGATAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2307
 Db 574 AsnSerProTyrGluAsnThrAsp-----LeuSer 583
 QY 2308 TTACAGTTAGGAGCTAAAGCCC-----CCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361
 Db 584 ArgAlaLeuTyrAlaGlnProMetLeuAlaIleSerGluAlaSerAspAsnGlnLeuGln 603
 QY 2362 AATGAG-----ATGCTAAGTATGCTTATCAAGGAAGCTGG 2397
 Db 604 SerGluSerMetAspPheSerLysValAsnValProHisTyrGlyTyrGlnGlyLeuTyr 623
 QY 2398 AAGCTGCGTGG-----CAT 2412
 Db 624 ThrTyrGlyTyrAlaLysThrGluAsnProThrThrThrProProAlaThrIleThrAsp 643
 QY 2413 CCTAATACAGCAATAATGCTTCTTATCTCTGAAAGCTACATGACTGCTAAAGCTGGGTAT 2472
 Db 644 ProLysLysAlaAsnGlnPheHisArgThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 663
 QY 2473 AATCCTGGGCTGAGCGAGTAGCTTCTTGTGTTCCAAATAGTTTATGCGGATCCATTTTA 2532
 Db 664 IleProSerProLysHisLysSerProLeuIleAlaAsnThrLeuTyrGlyAsnIleLeu 683
 QY 2533 GATATACGATCT-----GCCATTCAGCAATTCAGCAAGTGTGGTGGGCTCTTAT 2586
 Db 684 PheAlaThrGluAsnLeuLysAsnSerSerGlyGlnGluLeuLeuAspArgProPheTyr 703
 QY 2587 TGTCGAGGATTTGGGTTTCTGGAGTTTCGAAATTTCTTATCATGACCGCGATGCTTTA 2646
 Db 704 -----GlyIleThrGlyGlyGlyLeuGlyMetMetValTyrGlnGluProArgLysAsp 721
 QY 2647 GGTCCAGGATATCGGTATATTAGTGGGGTTATTTCTTCTTACGA-----GCAAACTCC 2697
 Db 722 HisProGlyPheHisMetHisThrSerGlyTyrSerAlaGlyMetIleThrGlyAsnThr 741
 QY 2698 TACTTTGATCATCGATGTTTGTCTAGCATTTACCAAGATTTTGGTAGA---TCTAAA 2754
 Db 742 HisThrPheSerLeuArgPheSerGlnSerTyrThrLysLeuAsnGluArgTyrAlaLys 761
 QY 2755 GATTATGATGTTGCTGTTTCCAAATCATCATCTGCTGATGATGCTGTTTATCTATCTAC 2814
 Db 762 AsnTyrVal-----SerSerLysAsnTyrSerCysGlnGlyGluMetLeuLeuSerLeu 779

QY	2815	CAACAAGCTTTATGTGATCCTATTGTTGCGAGATGCGTTTATCCGTGCTAGCTACGGG	2874	QY	682	GAGAACATACGG-----	ACTTCTACAAATGGGCGAGCTCTA	717
DB	780	GlnGluGlyLeuMetLeuThrLysLeuIleGly-----LeuTyrSer	793	DB	59	AspAsnLeuArgTyrIleLeuAlaIleLeuGlnLysThrProAsnGluGlyAlaAlaVal	78	
QY	2875	TTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAG-----AGCGATGTT	2928	QY	718	AGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTTAAAGAATTATCTTTTCC	777	
DB	794	TyrGlyAsnHisAsnSerHisHisPheTyrThrGlnGlyGluAspLeuSerSerGlnGly	813	DB	79	-----ThrIleThrAspTyrLeuSerPhePheAspThrGlnLysGluGlyIleTyrPhe	96	
QY	2929	CGTTGGGATAATAACTGCTGCTGCGAGAGATTTGGAGCGGATTAACCGATT-----	2979	QY	778	AATTGCAATTCATTACTTGGCGTACTGCTGCTGCAACGACTAATAAGGCTAGCCAGACT	837	
DB	814	GluPheHisSerGlnThrPheGlyGlyAlaValPhePheAspLeuProLeuLysProPhe	833	DB	97	AlaLysAsn-----LeuThrProGluSerGlyGlyAlaIleGlyTyrAlaSer	112	
QY	2980	-----GTGATTACTCCATCTAAGCTCTATTGTAATGAGTTGCGTCTTTTCGTG	3027	QY	838	CGGACGACCAACATCTACACCGCTCTAATGGTACTATTATTCTTAAACACAGATCTTTTGTA	897	
DB	834	GlyArgThrHisIleLeuThr-----AlaProPheLeu	844	DB	113	ProAsnSerProThrValGluIleArgAspThrIle-----GlyProValIlePhe	129	
QY	3028	CAAGCTGAGTTTCTTATGCGCATCATGAATCTTTTACAGAGGAGGCGATCAAGCTCGG	3087	QY	898	CTCAATAATAG-----AAGTTCTCTATTCTATAGTAATTTAGTCTCTGGAGAT-----	945	
DB	845	GlyAlaIleGlyMetTyrSerLysLeuSerSerPheThrGluValGlyAlaTyrProArg	864	DB	130	GluAsnAsnThrCysCysArgLeuPheThrTyrArgAsnProTyrAlaAlaAspLysIle	149	
QY	3088	GCATTC---AAGAGCGGACATCTCTAAATCTATCATCTGCTTGTGGAGTGAAGTTGAT	3144	QY	946	-----GGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGT	999	
DB	865	ThrPheIleThrGluThrProLeuIleAsnValLeuIleProIleGlyValLysGlySer	884	DB	150	ArgGluGlyGlyAlaIleHisAlaGlnAsnLeuTyrIleAsnHisAsnHisAspValVal	169	
QY	3145	CGATGTTCTAGTACATCTCTAAATCTATCATCTGCTTGTGGAGTGAAGTTGAT	3201	QY	1000	GTCTTCCAAAGAAATACTGCTCAAGCTGATGGGAGGCTTCTCAAGTAGTACCAGGTTTC	1059	
DB	885	PheMetAsnAlaThrHisArgProGlnAlaTyrThrValGluLeuAlaTyrGlnProVal	904	DB	170	GlyPheMetLysAsnPheSerTyrValGlnGlyGlyAlaIleSerThrAlaAsnThrPhe	189	
QY	3202	GCTTATCGCACCATCTCTGCTACTGAGCAACGCTCTATCCATCAAGAGACATGGACA	3261	QY	1060	TCTGCTATGGCTTAACAGAGGCTCTTATTCCTTTGTAGCGAATGTT-----	1104	
DB	905	LeuTyrArgGlnGluProSerIleSerThrGlnLeuLeuAlaGlyLysGlyMetTyrPhe	924	DB	190	ValValSerGluAsnGlnSerCysPheLeuPheMetAspAsnIleCysIleGlnThrAsn	209	
QY	3262	ACAGATCGCTTTTCATTTAGCAAGACATGAGTTGTGTTAGAGGATCTATGATGCTTCT	3321	QY	1105	-----GCAGGAGTAGAGGGGAGGATGCTGCTGTTTCAGGATGGCGACGAGGTGCA	1161	
DB	925	GlyHisGlySerProAlaSerArgHisAlaLeuAlaTyrLysIleSerGlnLysThrGln	944	DB	210	ThrAlaGly---LysGlyGlyAlaIleTyrAla-----GlyThrSer	222	
QY	3322	CTAACA-----AGTAATATAGAGTATATGCCATCGCAAGATAT	3360	QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCGGTAGAGTTT	1221	
DB	945	LeuLeuArgPheAlaThrLeuGlnLeuGlnTyrHisGlyTyrTyr	959	DB	223	AsnSerPheGlu-----SerAsnAsnCysAspLeuPhePhe	234	
RESULT 36				QY	1222	GATGGGAACGTAGCCGAGTAGGAGGAGATTACTCC-----	1260	
E71460				DB	235	IleAsnAsnAlaCysCysAlaGlyGlyAlaIlePheSerProIleCysSerLeuThrGly	254	
probable outer membrane protein E - Chlamydia trachomatis (serotype D, strain UW3/Cx)				QY	1261	---TAGCGNACGTTGCTTCTGATTAATGGAACCTTGTCTTCAACAATGTTGCT	1317	
C;Species: Chlamydia trachomatis				DB	255	AsnArgGlyAsnIleValPheTyrAsnAsn-----ArgCysPheLysAsn-----	269	
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004				QY	1318	TCTCCTGTTTACATTGCTGCTAAGCAACCAACAGTGGACAGCGCTTCTAATACGAGTAAT	1377	
C;Accession: E71460				DB	270	-----ValGluThrAlaSerSer	275	
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998				QY	1378	AATTACGAGATGGAGGAGCTATCTTCTGTAAGAATGGTGGCCAGCAGGATCCAATAAC	1437	
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis				DB	276	GluAlaSerAspGlyGlyAlaIleLysValThrArgLeuAspValThrGlyAsn---	294	
A;Reference number: A71570; MUID:9900809; PMID:9784136				QY	1438	TCTGATCAGTTCTTCTTGTAGGAGGAGTAGTTTCTTTTAGTAGCAATGATGCTGCT	1497	
A;Accession: E71460				DB	295	-----ArgGlyArgIlePhePheSerAspAsnIleThrLys	306	
A;Status: preliminary				QY	1498	GGGNAAGGGGAGCTATTTATGCCAAAAGCTCTCGTTGCTAAGTGTGCCCTGTACAA	1557	
A;Molecule type: DNA				DB	307	AsnTyrGlyGlyAlaIleTyrAlaProValValThrLeuValAspAsnGlyProThrTyr	326	
A;Residues: 1-964 <ARN>				QY	1558	TTTTTAAGGAATATCGCTAATGAT---GGTGGAGCGGATTTATTAGGAGATCTGGAGAG	1614	
A;Cross-references: UNIPROT:084877; UNIPARC:UPI0000131CF3; GB:AE001360; GB:AE001273; NID				DB	327	PheIleAsnAsnIleAlaAsnAsnLysGlyGlyAlaIleTyrIleAspGlyThrSerAsn	346	
A;Experimental source: serotype D, strain UW-3/Cx				QY	1615	CTCAGTTTATCTGCTGATTTATGGAGATATTATTTTCGATGGGAATCTTAAAGAACAGCC	1674	
A;Genetics: pmpE				DB	347	SerLysIleSerAlaAspArgHisAlaIlePheAsnGluAsnIle-----	362	
Alignment Scores:								
Pred. No.:								
Score:								
Percent Similarity:								
Best Local Similarity:								
Query Match:								
DB:								
US-10-701-844-1 (1-4435) x E71460 (1-964)								
QY	631	AACTTATTAGGAGTTTACTGTTTATGGGAGAGACACTCGTTGACT-----TTC	681	QY	631	AACTTATTAGGAGTTTACTGTTTATGGGAGAGACACTCGTTGACT-----TTC	681	
DB	39	SerLeuSerAsnLysIleSerLeuThrGlyAspThrHisAsnLeuThrAsnCysTyrLeu	58	DB	39	SerLeuSerAsnLysIleSerLeuThrGlyAspThrHisAsnLeuThrAsnCysTyrLeu	58	

QY 1675 AAAGAGATGTCGCGATGTTAATGGCGTAACGTGTCTCTCA-----CAA 1719
 Db ValThrAsnValThrAsnAlaAsnGlyThrSerThrSerAlaAsnProProArgAsn 382
 QY 1720 GCATTTTCGATGGGATGGGAGAAATAACGACATTAAAGCATTAAGCGAGGCGATCAG 1779
 Db AlaIleThrValAlaSerSerSerGlyGluLeuLeuGlyAlaGlySerGlnAsn 402
 QY 1780 ATTCTCTTAATGATCCCATCGAGATGGCAACGGAATAACACCGCAGCGAGCTCTCC 1839
 Db LeuIlePheTyrAspProIleGluValSerAsn-----AlaGly 415
 QY 1840 AAATCTCTAAATAATTAAAGATGCTGAAGATACACAGGGGATATTGTTTTCCTAAT--- 1896
 Db ValSerValSerPheAsnLysGluAlaAspGlnThrGlySerValValPheSerGlyAla 435
 QY 1897 -----GGAAGCAGTACTTTGTACCAAAAT-----GTT 1923
 Db ThrValAsnSerAlaAspPheHieGlnArgAsnLeuGlnThrLysThrProAlaProLeu 455
 QY 1924 AGATAGAGCAAGAGGATGTTCTTCGTGAAAGCGCAAAATATCATGTAATCTCTTA 1983
 Db ThrLeuSerAsnGlyPheLeuCysIleGluAspHieAlaGlnLeuThrValAsnArgPhe 475
 QY 1984 AGTCAGACGGTGGGAGTCTGTATATGAGAGCTGGAGT----- 2022
 Db ThrGlnThrGlyGlyValValSerLeuGlyAsnGlyAlaValLeuSerCysTyrLysAsn 495
 QY 2023 ---ACATGGGATTTTGTAATCTCCACACCAACACAGCCCTCCGCGCTAATCAGTTG 2079
 Db GlyThrGlyAsp-----SerAlaSerAsnAlaSer 505
 QY 2080 ATCAGCTTTCCAATCTGCATTTGCTCTTCTCTTTGTTAGCAACAATGCA----- 2133
 Db IleThrLeuLysHieIleGlyLeuAsnLeuSerSerIleLeuLysSerGlyAlaGluIle 525
 QY 2134 -----GTTACGAATCCTCTACCAATCCTCCAGCGCAAGATTCATCCTCGCATCT 2187
 Db ProLeuLeuTrpValGluProThrAsn-----AsnSerAsnAsnTyrThrAlaAspThr 543
 QY 2188 GGTAGCACA---ACTGCTGTTCTGTACAAATTAGTGGCGCTATCTTTTTCAGCATTTG 2244
 Db AlaAlaThrPheSerLeuSerAspValLysLeuSer-----LeuIleAspAspTyr 560
 QY 2245 GATGATACAGCTATGATAGGTATGATGGCTAGTCTTCTTAATCAAAAAATCAATGCTCG 2304
 Db GlyAsnSerProTyrGluSerThrAsp-----Leu 570
 QY 2305 AAATTACAGTTAGGACTAAGCCC-----CCAGCTAATGCCCATCAGATTTGACTCTA 2358
 Db ThrHieAlaLeuSerSerGlnProMetLeuSerIleSerGluAlaSerAspAsnGlnLeu 590
 QY 2359 GGGAAATGAG-----ATGCCCTAAGTATGGCTATCAAGGAAGC 2394
 Db GlnSerGluAsnIleAspPheSerGlyLeuAsnValProHieTyrGlyTrpGlnGlyLeu 610
 QY 2395 TCGAAGCTTTCGTGG----- 2409
 Db TrpThrTrpGlyTrpAlaLysThrGlnAspProGluProAlaSerSerAlaThrIleThr 630
 QY 2410 GATCCTAATACAGCAATAATGGTCTTATCTCTGAAAGCTACATGGACTAAAGCTGGG 2469
 Db AspProGlnLysAlaAsnArgPheHieArgThrLeuLeuLeuLeuThrTrpLeuProAlaGly 650
 QY 2470 TATAATCCTGGCGCTGAGCGAGTAGCTCTTTGTTCCAAATAGTTATGGGGATCC--- 2526
 Db TyrValProSerProLysHieArgSerProLeuIleAlaAsnThrLeuTrpGlyAsnMet 670
 QY 2527 ATTTTAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTAT 2586
 Db LeuLeuAlaThrGluSerLeuLysAsnSerAlaGluLeuThrProSerGlyHieProPhe 690
 QY 2587 TGTCGAGGATATGGGTTTCTGGAGTTTCGAATTTCTTCTATCATGACCGCGATGCTTTA 2646

Db 691 Trp---GlyIleThrGlyGlyGlyLeuGlyMetMetValTyrGlnAspProArgGluAsn 709
 QY 2647 GGTGAGGATATCGGTATATATTAGTGGGGTTATTCTCTTAGCA-----GCAAACTCC 2697
 Db HisProGlyPheHieMetArgSerSerGlyTyrSerAlaGlyMetIleAlaGlyGlnThr 729
 QY 2698 TACTTTGGATCATCGATGTTTGGTCTAGCATATTACCGAAGTATTGTTAGTAGA---TCTAAA 2754
 Db HisThrPheSerLeuLysPheSerGlnThrTyrThrLysLeuAsnGluArgTyrAlaLys 749
 QY 2755 GATTATGTAGTGTCTGTTCCAAATCATCTGTCATAGGATCCGTTTCTCTATCTATCTACC 2814
 Db AsnAsnVal-----SerSerLysAsnTyrSerCysGlnGlyGluMetLeuPheSerLeu 767
 QY 2815 CAACAAGCT-----TTATGTGGATCTCTATTGTTTCGGAGATGCGGTTTATC 2859
 Db GlnGluGlyPheLeuLeuThrLysLeuValGlyLeuTyrSerTyrGlyAspHieAsnCys 787
 QY 2860 CGTCTAGCTACGGGTTTGGGAATCAGCATATGAACCTCATATACATTTTCAGAGGAG 2919
 Db HisHisPheTyrThrGlnGly---GluAsnLeuThrSerGlnGlyThrPheArgSerGln 806
 QY 2920 AGCGATGTTGTTGGGATAATAACTGCTGCTCGAGAGATTGGAGCGGATTACCGATT 2979
 Db Thr-----MetGlyGlyAlaValPhePheAspLeuProMet 818
 QY 2980 GTGATTACTCATCTAAGCTCTATTGTAATGAGTTGCGTCTTTGTCGCAAGCTGAGTTT 3039
 Db LysProPheGlySerThrHieIleLeuThr-----AlaProPheLeuGlyAlaLeuGly 836
 QY 3040 TCTTATGCGCATCATGAATCTTTTACAGAGAGGCGATCAAGCTCGGCGATTCAAGAGC 3099
 Db IleTyrSerSerLeuSerHiePheThrGluValGlyAlaTyrProArgSerPheSerThr 856
 QY 3100 ---GGACATCTCTAAATCTATCATCTCTCTGAGTGAAGTTTGATCGATGTTCTAGT 3156
 Db LysThrProLeuIleAsnValLeuValProIleGlyValLysGlySerPheMetAsnAla 876
 QY 3157 ACACAT---CCTAATAATATAGCTTTTATGCGGCTTATATCTGTGATGCTTATCGCAC 3213
 Db ThrHieArgProGlnAlaTyrThrValGluLeuAlaTyrGlnProValLeuTyrArgGln 896
 QY 3214 ATCTCTGTACTAGACAAACGCTCTATCCATCAAGACATGAGACAGATGCGCTTT 3273
 Db GluProGlyIleAlaAlaGlnLeuLeuAlaSerLysGlyIleTyrPheGlySerGlySer 916
 QY 3274 CATTTAGCAACATGAGTTGTGTTAGAGATCTATGTATGCTTCT----- 3321
 Db ProSerSerArgHieAlaMetSerTyrLysIleSerGlnGlnThrGlnProLeuSerTrp 936
 QY 3322 CTAAACAAGTAATAGAAAGTATATGCG 3348
 Db LeuThrLeuHiePheGlnTyrHieGly 945
 RESULT 37
 F86548
 polypeptide outer membrane protein B family [imported] - Chlamydomonadales pneumoniae (strain
 C:Species: Chlamydomonadales pneumoniae, Chlamydomonadales pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: F86548
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishi-
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; PMID:20330349; PMID:10871362
 A:Accession: F86548
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-938 <STO>
 A:Cross-references: UNIPROT:Q92883; UNIPARC:UPI000004707C; GB:BA000008; NID:98978836; PI
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: pmp_15

Alignment Scores: 2,74e-36 Length: 938
 Pred. No.: 644.00 Matches: 286
 Score: 41.3% Conservative: 160
 Best Local Similarity: 25.8% Mismatches: 408
 Query Match: 8.2% Indels: 198
 DB: 44 Gaps: 44

US-10-701-844-1 (1-4435) x F86548 (1-938)

```

QY 622 TGTGTTGGAACTATTAGG---AGTTTACTGTTTATAGGAGAGACACTCGTTGACT 678
DB 5 CysPheGlyMetLeuLeuProPheThrPheValLeuAlaAsnGlyLeuGlnLeuPro 24
QY 679 TTCGAGAACATCGGACTCTTACAAATGGGCGAGCTTAAGT-----AATAGCGCT 729
DB 25 LeuGluThrTyrlleThr-----LeuSerProGluTyrlleGlnAla 38
QY 730 GCTGATGGAGTGTACTATTGAGGTTTTAAAGAAATTATCCTTT---TCCAAATGCAAT 786
DB 39 ProGlnValGlyPheThrHisAsnGlnAsnGlnAsnLeuAlaValGlyAsnHisAsn 58
QY 787 TCATTACTTCCGCTACTGCTGCAACGACTAATAAGGAGTACCGAGCT-----837
DB 59 AppPheIleLeuAspTyrlsTyrlsTyrlsTyrlsTyrlsTyrlsTyrlsTyrls 78
QY 838 -----CCGACGACCAATCTACACCGTCT 861
DB 79 LeuLeuIleSerGluAsnIleGlyAsnValPhePheGluLysAsnValCysProAsnSer 98
QY 862 AATGGTACTATTATTCTAAACAGATCTTTGTTACTCAATAGAGAGTTCTCATTC 921
DB 99 GlyGlyAlaIleTyrlsAlaGlnAsnCysThrIleSerLysAsnGlnAsnTyrlsAlaPhe 118
QY 922 TATAGTAATTTAGTCTCTGGAGAT-----GGGGGAGCT 954
DB 119 ThrThrAsnLeuValSerAspAsnProThrAlaThrAlaGlySerLeuLeuGlyGlyAla 138
QY 955 ATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGTTCTCCAGAAAT 1014
DB 139 LeuPheAlaIleAsnCysSerIleThrAsnAsnLeuGlyGlnGlyThrPheValAspAsn 158
QY 1015 ACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACCAAGTTTCTCTGCTATGGCTAAC 1074
DB 159 LeuAlaLeuAsnLysGlyGlyAlaLeuTyrlsThrGluThrAsnLeuSerIleLysAspAsn 178
QY 1075 GAGGCTCTATTGCTTTGTAGCGAATGTTGCA-----GGAGTAAGAGGGGAGGG 1125
DB 179 LysGlyProIleIleIleLysGlnAsnArgAlaLeuAsnSerAspSerLeuGlyGly 198
QY 1126 ATTGCTGCTGTTTCCAGGATGGGACGAGGAGTGTCAATCTACTTCAACAGAAATGCCA 1185
DB 199 IleTyrlsSer-----GlyAsnSerLeuAsnIleGlu-----208
QY 1186 GTAGTAGTTTTCAGAAATATCTGCGTAGAGTTTCAATGGGACGTAGCCCGAGTAGGA 1245
DB 209 -----GlyAsnSerGlyAlaIleGlnIleThrSerAsnSerSerGlySerGly 224
QY 1246 GGAGGATTTACTCTACGGGAAGTGTGTTCTCTGAAATATGAAACAACTTGTGTTCTC 1305
DB 225 GlyGlyIlePheSerThrGlnThrLeuThrIleSerSerAsnLysLysLeuIleGluIle 244
QY 1306 AACAAATGTTGCTTCTCTGTTTACATTTGCTGCTAAGCAACCAAGTGGACAGGCTTCT 1365
DB 245 SerGlu-----AsnSerAlaPheAlaAsn 252
QY 1366 AATACGAGTAAATATTCGAGATGAGGA-----GCTATCTTCTGTAGAAAT 1413
DB 253 AsnTyrlsSerAsnPheAsnProGlyGlyGlyGlyLeuThrThrPheCys-----270
QY 1414 GTGCGGCAAGCAGGATCCAACTACTCGGATCGATTTCTTCTTGTGATGAGGAGGTAGTT 1473
DB 1414 GTGCGGCAAGCAGGATCCAACTACTCGGATCGATTTCTTCTTGTGATGAGGAGGTAGTT 1473

```

```

DB 271 -----ThrIleLeuAsnAsnArgGluGlyValLeu 280
QY 1474 TTCTTTAGTACAAATAGTCTGGGAAAGGGGAGCTATTATTATGCAAAAGACTCTCG 1533
DB 281 ----PheAsnAsnAsnGlnSerGlnSerAsnGlyGlyAlaIleHisAlaLysSerIle 299
QY 1534 GTTGCTAACTGTGGCCCTGTCAATTTTAAAGGAATATCGCTAATGATGTGGAGCGATT 1593
DB 300 IleLysGluAsnGlyProValTyrlsPheLeuAsnAsnThrAlaThrArgGlyGlyAlaLeu 319
QY 1594 TATTTA-----GGAGAACTCTGGAGAGCTCAGTTTATCTCTGATTATGAGAT 1641
DB 320 LeuAsnLeuSerAlaGlySerGlyAsnGlySerPheIleLeuSerAlaAsnGlyAsp 339
QY 1642 ATTTATTTTCGATGGAATCTTAAAGAAAGACAGCAAGAGAGATGTCGCCGAT-----GTT 1695
DB 340 IleIlePheAsnAsnAsn-----ThrAlaSerLysHisAlaLeuAsnProProTyrls 356
QY 1696 AATGCGGTAACTGTCTCTCAAGCCATTTTCGATGGGATCGGAGGGGAAATAACGACA 1755
DB 357 ArgAsnAlaIleHisSerThrProAsnMetAsnLeuGlnIleGly-----371
QY 1756 TTAGAGCTAAAGCAGGCGCATCAGATTCTCTTAAATGATCCCATCGAGATGCAACGGA 1815
DB 372 -----AlaArgProGlyTyrlsArgValLeuPheTyrlsAspProIleGlu-----385
QY 1816 AATAACCGCCAGCGCAGTCTTCCAAACTCTTAAAAAATTAAACGATGGTGAAGATACACA 1875
DB 386 -----HisGluLeuProSerSerPheProIleLeuPheAsnPheGluThrGlyHisThr 403
QY 1876 GGGGATATTGTTTGTAAAGAGCAGTACTTTGTACAAAATGTTACG-----1926
DB 404 GlyThrValLeuPheSer-----GlyGluHisValHisGlnAsnPheThrAspGluMet 421
QY 1927 -----ATAGACAGGAGGAGTGTGTTCTCT 1953
DB 422 AsnPheSerSerTyrlsLeuArgAsnThrSerGluLeuArgGlnGlyValLeuAlaValGlu 441
QY 1954 GAAAGGCAAAATATCAGTGAATTTCTTAAGTCAGACAGGTCGGAGTCTGTATATGGA 2013
DB 442 AspGlyAlaGlyLeuAlaCysTyrlsPhePheGlnArgGlyGlyThrLeuLeuGly 461
QY 2014 GCTGGGAGT-----ACATGGGATTTTGTAACTCCACCAACCAACGAGCTCTCGCC 2067
DB 462 GlnGlyAlaValIleThrAlaGlyThrIleProThrProSerSerThrProThrThr 481
QY 2068 GCTAATCAGTTGATCAGCTTCCAAATCGCATTTGTCTCTTCTCTTCTTTGTAGCAAC 2127
DB 482 ValGlySerThrIleThrLeuAsnHisIleAlaIleAspLeuProSerIleLeuSerPhe 501
QY 2128 AATGAGTTACGATCCTCTACCAATCCTCCAGCGCAAGATTTCTCATCTGCGATCAAT 2187
DB 502 GlnAla-----GlnAlaProLysIleTyrlsIleTyrlsProThrLysThr 515
QY 2188 GGTAGCACCA-----ACTGCTGTTTCTGTACAAATTAGTGGGCTATCTTTTTT 2235
DB 516 GlySerThrTyrlsThrGluAspSerAsnProThrIleThrIleSerGlyThrLeuThrLeu 535
QY 2236 GAGGATTTG---GATGATACAGCTTATGATAGTATGATGTAGTGTCTTATCAATAAAA 2292
DB 536 ArgAsnSerAsnAsnGluAspProTyrlsAspSerLeuAspLeuSerHisSerLeuGlyLys 555
QY 2293 ATCAATGCTCTGAAATTA---CAGTTAGGAGCTAAGCCCCCAGCTAATCCCCATCAGAT 2349
DB 556 ValProLeuLeuTyrlsIleValAspValAlaAlaGlnLysIleAsnSerSerGlnLeuAsp 575
QY 2350 TTGACTCTAGGGAATCAGATGCTTAAGTATGCTATCAAGAGAGCTGGAGCTTGGTGG 2409
DB 576 LeuSerThrLeuAsnSerGlyGluHisTyrlsGlyTyrlsGlnGlyIleTyrlsThrTyrls 595
QY 2410 GATCCTTAATACA-----GCAAAATATGTCCTTAT 2439
DB 596 ValGluThrThrThrIleThrAsnProThrSerLeuLeuGlyAlaAsnThrLysHisLys 615

```



```

Db      815 LeuAlaAlaPheGlnGluSerGlyAspHisAlaArgGluPhe---SerLeuHisArgPro 833
Qy      3106 CTCCTAAATCATCATGTTCTCTGTTGGAGTGAAGTTTGTATCGATGCTTCAGTACACATCCT 3165
Db      834 LeuThrAspValSerLeuProValGlyLe---ArgAlaSerTyrLeuAsnHis 850
Qy      3166 AATAAATATAGCTTTATG-----CGGGCTTATATCTGTGATGCTTATTCGCACC 3213
Db      851 HisArgValProLeuValTyrLeuThrGluLeuSerTyrArgSerThrLeuTyrArgGln 870
Qy      3214 ATCTCTGCTACTGAGACACAGCTCTCATCCATCAAGACATGACACACAGATGCTCTTT 3273
Db      871 AspProGluLeuHiserLysLeuLeuLeuSerGlnGlyThrTrpThrGlnAlaThr 890
Qy      3274 CATTTAGCAACACAT-----GGAGTGTGTTAGGATCATG-----TATGCTTCT 3321
Db      891 ProValThrTyrAsnAlaLeuGlyTyrLeuValysAsnThrMetGlnValPheProlys 910
Qy      3322 CTAAACAAGTAATAGAAATATATGCGCATGGAAGATATGATGATCGAGATGCTTCTCGA 3381
Db      911 ValThrLeuSerLeuAspTyrSerAla-----AspIleSerSerThrLeu 926
Qy      3382 GCCTATGCTTTGAGTGCAGGAAGTATGACTCCGGTTC 3417
Db      927 SerHisTyrLeuAsnValAlaSerArgMetArgPhe 938

RESULT 39
A:1731
polymorphic membrane protein D family TC0197 [imported] - Chlamydia muridarum (strain N1)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81731
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81731
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1520 <TET>
A:Cross-references: UNIPROT:Q9PLB0; UNIPARC:UPI0000057825; GB:AE002160; NID
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0197

Alignment Scores:
Pred. No.: 5,728-34 Length: 1520
Score: 611.00 Matches: 291
Percent Similarity: 37.9% Conservative: 164
Best Local Similarity: 24.2% Mismatches: 454
Query Match: 7.8% Indels: 291
DB: 2 Gaps: 55

US-10-701-844-1 (1-4435) x A81731 (1-1520)
Qy      405 CTTTCTTCAATGATTCT---AGCTTATTTCTGTGCTCTTTAAATGGGGGGGATATGC 461
Db      417 IleSerPheGluGlySerLysSerPheGlyGlyAlaIleAlaCysGlyAsnPheSer 436
Qy      462 AGCAAAATCATGTTTCC-----TCAGAAATTTACGATGGGGA 500
Db      437 SerGluAsnAsnSerSerAlaLeuGlySerIleAspIleSerAsnAsnLeu-----Gly 454
Qy      501 GAGGTTAACTGTATCATTTCTCCCTATCTGTTATAGGATCCGAGTGGGACTCTGTTT 560
Db      455 AspIleSerPheLeuArgThrLeu-CysThrThrSerAspLeuGlyGlnThrAspTyrGl 474
Qy      561 TTTCTGCAGAGAGATTAAACATTAAATAATCTTGACAAATCTATTGCAAGCTTTGCTCTTAAG 620
Db      474 nGlyGlyAlaLeuPheAlaGluAsnIleSerLeuSerGluAsnAlaGlyAlaIleTh 494

```

```

Qy      621 TTGTTTGGGAATTATTAGGAGTTTACT-----GTTTGGGAGGAGACA 668
Db      494 rPheLysAspAsnIleValLysThrPheAlaSerAsnGlyLysMetLeuGlyGlyAl 514
Qy      669 CTCCTGACTTTCGAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGT----- 720
Db      514 alleLeuAlaSerGlyAsnValLeuIleSerLysAsnSerGlyGluLeuSerPheValGl 534
Qy      721 -AATAGCGCTGCTGATCGACTGTTTACTATTATGAGGGTTTAAAGAATATCTTTTCAA 779
Db      534 YAsnAlaArgAlaProGlnAlaIleProThrArgSerSerAspGluLeuSerPhe----- 552
Qy      780 TTGCAATTCATTACTCCGCTACTGCTGCTGCAACGACTAATAAGGTAGCCAG---AC 836
Db      553 -----GlyAlaGlnLeuTh 557
Qy      837 TCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTAAACACAGACTTTTGT 896
Db      557 rGlnThrThrSerGlyCysSerGlyGlyAlaLeuPheGlyLys---GluValAlaIle 576
Qy      897 ACTCAATPAATGAGAAGTTCTCATTTCTATAGTAATTTAGTCTCT----- 939
Db      576 eValGlnAsnAlaThrValPheGluGlnAsnArgLeuGlnCysGlyGluGlnGluTh 596
Qy      940 ----GGAGATGGGGAGCTATA---GATGCTAAGAGCTTAACGTTCAAGGAATTAGCAA 992
Db      596 rHisGlyGlyGlyAlaValTyrGlyMetGluSerAlaSerIleIleGly---AsnSe 615
Qy      993 GCTTTGTGCTTCCCAAGAAATATCTGCT-----CAAGCTCATGGGGAGCTTTGTCA 1043
Db      615 rPheValArgPheGlyAsnAsnTyrAlaValGlyAsnGlnIleSerGlyGlyAla----- 633
Qy      1044 AGTAGTCACCAAGTTTCTGCTATGCTATGCTTAACAGAGCTCTATTGCTTGTAGGAAATGT 1103
Db      634 -LeuLeuSerLysLysValArgLeuAlaGluAsnThrArgValAspPheSerArgAsnIle 653
Qy      1104 TGCAGGAGTAGAGGGGGGGGATGCTGCTGTTTCAAGATGGGCGAGCAGGAGTGTCTATC 1163
Db      653 eAlaThrPheCysGlyGlyAlaVal---GlnValSerAspGly-----Se 667
Qy      1164 ATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGGTAGAGTTTGA 1223
Db      667 rCysGluLeuIleAsnAsnGlyTyrValLeuPheArgAspAsnArgGlyGlnThrPheGl 687
Qy      1224 TGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCAGCGGAAC-----GTTGC 1274
Db      687 yGlyAlaIleSerCysLeuLysGlyAspValIleIleSerGlyAsnLysAspArgValGl 707
Qy      1275 TTTCTCGAATTAAT----- 1287
Db      707 uPheArgAspAsnIleValThrArgProTyrPheGluGluAsnGluGluLysValGluTh 727
Qy      1287 ----- 1287
Db      727 rAlaAspIleAsnSerAspLysGlnGluAlaGluGluArgSerLeuLeuGluAsnIleGl 747
Qy      1288 -----GGAAAAACCTTGTTC-----AACATGTTGCTTC 1319
Db      747 uGlnSerPheIleThrAlaThrAsnGlnThrPhePheLeuGluGluGluLysLeuProse 767
Qy      1320 TCCTGTTTACATTCCTGCTAAGCAACCAACCAAGTGGACAG----- 1359
Db      767 rGluAlaPheIleSerAlaGluGluLeuSerLysLysArgGluCysAlaGlyAlaIle 787
Qy      1360 -----GCTTCTAATACG-----AGTAATAATTA 1382
Db      787 ePheAlaLysArgValTyrIleThrAspAsnLysGluProIleLeuPheSerHisAsnPh 807
Qy      1383 CGGAGAT-----GGAGAGCTATCTTCTGTAAGAATGGTGGCAGCAGATCAATAA 1436
Db      807 eSerAspValTyrGlyGlyAlaIlePhe-----ThrGlySerLeuGlnGluThrAspLy 825
Qy      1437 CTCTGGATCAGTTTCTTTGAT-----GGAGAGGAGTAGTGTTCCTTTAG 1481

```

```
Db      825 sGlnaspValValThrProGluValValIleSerGlyAenAaspGlyAaspValIlePheSe 845
      QY      1482 TAGCAATGTAGCTGCGGAAA-----GGGGAGCTATTATGCG 1520
      Db      845 rGlyAenAlaAlaLysHisAaspLysHisLeuProAaspThrGlyGlyAlaIleCysTh 865
      QY      1521 CAAGAAGCTCTGGTGGCT--AACTGTGCGCCGTGTACAAATTTTAAAGGAATATCGCTAA 1577
      Db      865 rGlnAenLeuThrIleSerGlnAenAaspGlyAenValLeuPheLeuAenAaspPheAlaCy 885
      QY      1578 TGATGTGAGCGGCAATTTATAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATATGCG 1637
      Db      885 sSerGlyGlyAlaValArgIleGluAaspHisGlyGluValLeuLeuGluAlaPheGlyGl 905
      QY      1638 AGATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTGCCGATTTAA 1697
      Db      905 yAaspIleIlePheAenGlyAenSerPheArgAlaGln----- 918
      QY      1698 TGGCGTAACTGTCTCTCAAGCCATTTTCGATG---GGATCGGAGGGAATAACGAC 1754
      Db      919 -----GlySerAaspAlaIleTyrPheAlaGlyLysAaspSerArgIleLysAl 934
      QY      1755 ATTAAGAGCTAAAGCAGGCGATCAGATTCTCTTAATGATCCATCGAGATGCCAAACGG 1814
      Db      934 aLeuAenAlaThrGluGlyHisAlaIleValPheGlnAaspAlaLeuValPheGlu----- 952
      QY      1815 AAATAACCAAGCCAGCGAGCTTTCCAAACTTCTTAAATAATTAAC-----GATGGTGAAGG 1868
      Db      953 -AenIleGluGluArgLysSerSerGlyLeuLeuValIleAenSerGlnGluAenGluGl 972
      QY      1869 ATACACAGGGATATGTTTCTTAATGAGACAGTACTTTGTACCAAAATCTTACGAT 1928
      Db      972 yThrThrGlySerValArgPheLeuGlySerGlySerLysValProGlnTrpIleHisVa 992
      QY      1929 AGAGCAAGGA----- 1938
      Db      992 iGlnGlnGlyLeuLeuLeuHisGlyAlaIleLeuCysSerTyrcGlyValLysGl 1012
      QY      1939 -----AGGATGTTCTCTCGTGAAGGCAAAATATCATGTAATCTCTTAAG 1985
      Db      1012 nAaspProArgAlaLysIleValLeuSerAlaGlySerLysLeuLysAlaLeuAaspSerGl 1032
      QY      1986 TCAGACAGTGGAGCTGTATATGAGCTGGAGCTGACATGCGATTTGTAATCTCCACA 2045
      Db      1032 uGlnGluAenAenAla-----GluIleGlyAaspLeuGluAaspSerValAenSerGl 1049
      QY      2046 ACCACCA-----CAACAGCTCTCTCGCGCTAAATCAGTTGATCAGCT 2087
      Db      1049 uLysThrProSerLeuTrpIleGlyLysAenAlaGlnAlaLysValProLeuValAaspIl 1069
      QY      2088 TTCCAATCTGCATTTGCTCTCTTTCTTTGTAGCAACAACATGCAATGATCTCTCC 2147
      Db      1069 eHisThrIleSerIleAaspLeuAlaSer---PheSerSerLysAlaGlnGluThrPro-- 1087
      QY      2148 TACAATCTCCAGCGCAAGATCTCATCTGCGATGCTATGCTGAGCAACACTGCTGTTCC 2207
      Db      1088 -----GluGluAlaProGlnValIle-----ValProLysGlySe 1099
      QY      2208 TGTTACAATTAGTGGCGCTATCTTTTGTAGGATTTGAGTATGATACA-----GCTTATGA 2261
      Db      1099 rCysValHiserGlyGluLeuSerLeuGluLeuValAasnThrThrGlyLysGlyTyrgl 1119
      QY      2262 TAGTATGATTTGGCTAGTCTTCTAATCAAAAATCAATGTCCTG----- 2304
      Db      1119 uAenHisAlaLeuLeuLysAenAaspThrGlnValSerLeuMetSerPheLysGluGluAs 1139
      QY      2305 -----AAATTCAGTTAGGAGTAAAGCCAGCCAGCTAATGCG 2339
      Db      1139 nAaspGlySerLeuGluAaspLeuSerLysLeuSerVal----- 1151
      QY      2340 CCCATCAGATTTGACTCTAGGATGAGATGCTCT-----AAGTATGCGTA 2384
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 13, 2006, 11:45:09 ; Search time 116.9 Seconds
(without alignments)
4755.534 Million cell updates/sec

Title: US-10-701-844-1
Perfect score: 7883
Sequence: 1 gggcaaaactctcccccg.....gcattcctaagaaagaattc 4435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10701844/runat 12052006 165436 26532/app query.fasta_1
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs802p
-USER=US10701844 @CGN 1 1 307 @runat 12052006 165436 26532 -NCPU=6 -ICPU=3
-NO.MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=10 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	66.8	1012	4	US-10-701-844-2
2	5267	66.8	1012	4	US-10-766-711-2
3	5267	66.8	1012	5	US-10-931-779-2
4	5131.5	65.1	1013	4	US-10-467-534-9
5	5131.5	65.1	1013	5	US-10-498-327-5
6	5131.5	65.1	1013	5	US-10-498-327-81
7	5125.5	65.0	1013	5	US-10-498-327-83
8	5123.5	65.0	1013	5	US-10-498-327-9
9	5123.5	65.0	1013	5	US-10-498-327-13
10	5123.5	65.0	1013	5	US-10-498-327-93
11	5121.5	65.0	1013	5	US-10-498-327-1

12	5121.5	65.0	1013	5	US-10-498-327-3	Sequence 3, Appli
13	5116	64.9	984	4	US-10-701-844-43	Sequence 43, Appl
14	5116	64.9	984	4	US-10-766-711-43	Sequence 43, Appl
15	5116	64.9	984	5	US-10-931-779-43	Sequence 43, Appl
16	5112.5	64.9	1013	5	US-10-498-327-17	Sequence 17, Appl
17	5106.5	64.8	1013	4	US-10-701-844-15	Sequence 15, Appl
18	5106.5	64.8	1013	4	US-10-701-844-16	Sequence 16, Appl
19	5106.5	64.8	1013	4	US-10-766-711-15	Sequence 15, Appl
20	5106.5	64.8	1013	4	US-10-766-711-16	Sequence 16, Appl
21	5106.5	64.8	1013	5	US-10-931-779-15	Sequence 15, Appl
22	5106.5	64.8	1013	5	US-10-931-779-16	Sequence 16, Appl
23	5090	64.6	1006	3	US-09-841-132-190	Sequence 190, App
24	5090	64.6	1006	5	US-10-872-155-190	Sequence 190, App
25	5084	64.5	982	3	US-08-841-132-176	Sequence 176, App
26	5084	64.5	982	5	US-10-872-155-176	Sequence 176, App
27	3336.5	42.3	670	5	US-10-197-220-169	Sequence 169, App
28	3336.5	42.3	670	6	US-11-109-468-169	Sequence 17, Appl
29	2552	32.4	505	4	US-10-701-844-17	Sequence 17, Appl
30	2552	32.4	505	4	US-10-766-711-17	Sequence 17, Appl
31	2552	32.4	505	5	US-10-931-779-17	Sequence 17, Appl
32	2350	29.8	458	4	US-10-701-844-36	Sequence 36, Appl
33	2350	29.8	458	4	US-10-766-711-36	Sequence 36, Appl
34	2350	29.8	458	5	US-10-931-779-36	Sequence 36, Appl
35	1735	22.0	325	4	US-10-701-844-37	Sequence 37, Appl
36	1735	22.0	325	4	US-10-766-711-37	Sequence 37, Appl
37	1735	22.0	325	5	US-10-931-779-37	Sequence 37, Appl
38	1590.5	20.2	631	3	US-09-841-132-325	Sequence 325, App
39	1590.5	20.2	631	5	US-10-872-155-325	Sequence 325, App
40	1589	20.2	1016	3	US-09-841-280-95	Sequence 95, Appl
41	1589	20.2	1016	4	US-10-007-693-95	Sequence 95, Appl
42	1589	20.2	1016	4	US-10-467-534-41	Sequence 41, Appl
43	1589	20.2	1016	4	US-10-762-058-95	Sequence 95, Appl
44	1589	20.2	1016	5	US-10-197-220-95	Sequence 95, Appl
45	1589	20.2	1016	5	US-10-498-327-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1
US-10-701-844-2
; Sequence 2, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-2

Alignment Scores:				
Pred. No.:	0	Length:	1012	
Score:	5267.00	Matches:	1012	
Percent Similarity:	100.0%	Conservative:	0	
Best Local Similarity:	100.0%	Mismatches:	0	
Query Match:	66.8%	Indels:	0	
DB:	4	Gaps:	0	

US-10-701-844-1 (1-4435) x US-10-701-844-2 (1-1012)

QY 382 ATGCAACGCTCTTCATAAGTTCTTCTTCAATGATTCTAGCTTATTCTTGCTCTT 441
|||||

1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
442 TTTAAATGGGGGGATGATGACGAGAAATCATGTTCTCTCAAGAAATTTACGATGGGGAG 501
21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
502 ACGTTAACTGATCATTTCCCTACTACTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
562 TCTGCAGGAGGTAAACATTAATAAATCTTGACAAATCTATATGAGCTTTGCCCTTTAAGT 621
61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
622 TGTTTTGGGAACATTTAGGGAGTTTACTGTTTTAGGGAGGACACTGTTGACATTTTC 681
81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
682 GAGAACATACGAGCTCTACAAATGGGGCAGCTCTAAGTAAATAGCGCTGCTGATGGACTG 741
101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120
742 TTTACTATTGAGGTTTTAAAGAAATATCTTTTCCAAATTCATTCATTTACTTCCGCTA 801
121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
802 CTGCCTGCTCAACGACTAATAAGGTAGCCAGACTCCGACGACAAACATCTACACCGTCT 861
141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer 160
862 AATGGTACTATTATCTTAAACAGACTTTTTGTTTACTCAATAATGAGAAGTCTTCATTC 921
161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGACTTAACGGTTCAA 981
181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
982 GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGTCAAGTGGGGAGCTGT 1041
201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
1042 CAAGTAGTCACAGTTCTCTGCTATGCTAACGAGGCTCTATTGCTCTTGTAGCGAAT 1101
221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
1102 GTTGAGGAGTAAGAGGGGAGGATGCTGCTGCTCAGATGGGACAGGAGGTGCA 1161
241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
1162 TCATCTACTCCACAGAGATCCAGTAGTAGTTTTCAGAAATCTCCGCTAGAGTTT 1221
261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
1222 GATGGAAAGTAGCCCGAGTAGGAGGATTTTACTCTACGGGAAGCTTCTCTTCTCTG 1281
281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
1282 AATAATGGAACCACTGTTTCTCAACAATGTGCTTCTCTGTTTCAATGCTGCTAAG 1341
301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaLys 320
1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATATTACGAGATGGAGGACTATC 1401
321 GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIle 340
1402 TTTCTGTAAGAATGGTGGCAACGAGATCCAAATACTCGGATCAGTTTCTTTTATGGA 1461
341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly 360
1462 GAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGCTGGAAAGGGGAGCTATTATGCC 1521
361 GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyrAla 380

1522 AAAAAGCTCTCGTTGCTAACTGTGCCCCTGTCAAAATTTTAAAGGAATATCGCTAATGAT 1581
381 LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp 400
1582 GTTGGAGCGATTTATTAGGAGAACTCTGGAGACTCAGTTTATCTGCTGATTTATGGAGAT 1641
401 GlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAsp 420
1642 ATTATTTTCGATCGAATCTTAAAGAACACCCAAAGAGAAATGTCGCGATGTTAATGGC 1701
421 IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly 440
1702 GTAACTGTGTCTCACAGCCATTTCGATGGGATCGGAGGGGAAAATAACACATTTAAGA 1761
441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460
1762 GCTAAAGCAGGCGATCAGATTCCTTTTAATGATCCCATCGAGATGGCAACGGAAATAAC 1821
461 AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn 480
1822 CAGCCAGCGAGTCTTCCAAACTTCTAAAAATTAACCATGCTGTAAGGATACACAGGGGAT 1881
481 GlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGlyGlyTyrThrGlyAsp 500
1882 ATTGTTTTTCTAATGGAAGCAGTACTTTGTACAAAATGTTACGATAGACAAAGGAGG 1941
501 IleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIleGluGlnGlyArg 520
1942 ATTGTTCTTCTGTAAGGCGAAATTTATCATGTAATTTCTTAAGTCAAGAGTGGGAGT 2001
521 IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyLysSer 540
2002 CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACACACAGCCT 2061
541 LeuTyrMetGluAlaGlySerThrThrPheValThrProGlnProProGlnGlnPro 560
2062 CCTCCGCTAATCAGTTGATCACGCTTTCCAACTGCAATTTGCTCTTCTTCTTTGTTA 2121
561 ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeuLeu 580
2122 GCAAACTAATGAGTTCGATCCCTACCAATCTCCAGGCGCAAGATTCCTCATCTCGCA 2181
581 AlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisProAla 600
2182 GTCAATGTTAGCACCAACTGCTGTTCTGTACAATTAGTGGGCTCTCTTTTTTGAGGAT 2241
601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp 620
2242 TTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCTAATCAAAAATCAATGTC 2301
621 LeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsnVal 640
2302 CTGAATTTACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCCATCAGATTCAGCTCAGGG 2361
641 LeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGly 660
2362 AATGAGATGCCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGCTGGGATCCTTAATACA 2421
661 AsnGluMetProLysTyrGlyTyrGlnGlySerThrLysLysAlaThrThrLysThrGlyTyrAsnProGly 680
2422 GCAATAATGTCCTTATCTTACTCTGAAGCTCATGAGCTAAACCTGGGTATTAATCCTGGG 2481
681 AlaAsnAsnGlyProTyrThrLeuLysAlaThrThrLysThrGlyTyrAsnProGly 700
2482 CTGAGGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCCATTTTAGATATACGA 2541
701 ProGluArgValAlaSerLeuValProAsnSerLeuThrGlySerIleLeuAspIleArg 720
2542 TCTGCGCATTCAGCAATTCAGCAAGTGGAGTGGCGCTCTATTGTCGAGGATTTATGCG 2601
721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeuTyr 740

```

QY 2602 GTTCTCGAGTTTGGAAATTTCTTCTATCATGACCGCGATGCTTTAGGTCCAGGATATCGG 2661
DB |||||||
QY 741 ValSerGlyValSerAsnPhePheTyHisAspArgAspAlaLeuGlyGlnGlyTyArg 760
DB |||||||
QY 2662 TATATTAGTGGGGTTTATTCCTTAGGAGCAAACTCCTACTTTGGATCATCGATGTTGGT 2721
DB |||||||
QY 761 TyrIleSerGlyGlyTySerLeuGlyAlaAsnSerTyPheGlySerSerMetPheGly 780
DB |||||||
QY 2722 CTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATATAGTAGTGTGTTCCATCAT 2781
DB |||||||
QY 781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsnHis 800
DB |||||||
QY 2782 CATGCTTCATAGATCCGTTTATCTATCTACCAACAAGCTTTATGTGATCCTATTGG 2841
DB |||||||
QY 801 HisAlaCysIleGlySerValTyLeuSerThrGlnGlnAlaLeuGlyGlySerTyLeu 820
DB |||||||
QY 2842 TTCGGAGATCGGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAAACTCA 2901
DB |||||||
QY 821 PheGlyAspAlaPheIleArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThrSer 840
DB |||||||
QY 2902 TATACATTTGCAGAGAGAGCGATGTTGCTGGGATTAATACCTGCTGGCTGGAGAGATT 2961
DB |||||||
QY 841 TyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuAlaGlyGluIle 860
DB |||||||
QY 2962 GGAGCGGATTACCGATTGTGATTACTCCATCTAAGCTCTATTTGAATGAGTTGCGTCTCT 3021
DB |||||||
QY 861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyLeuAsnGlnLeuArgPro 880
DB |||||||
QY 3022 TTCGTGCCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAAGCGCATCAA 3081
DB |||||||
QY 881 PheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAspGln 900
DB |||||||
QY 3082 GCTCGGCGATTCAAGAGCGGACATCTCCTAAATCTATCAGTTCTGTTGGAGTGAAGTTT 3141
DB |||||||
QY 901 AlaArgAlaPheLysSerGlyHisLeuLeuAsnLeuSerValProValGlyValLysPhe 920
DB |||||||
QY 3142 GATCGATGTTCTAGTACACATCTTAATATATAGCTTTATGCGGCTTATATCTGTGAT 3201
DB |||||||
QY 921 AspArgCysSerSerThrHisProAsnLysTySerPheMetAlaAlaTyIleCysAsp 940
DB |||||||
QY 3202 GCTTATCGCACCATCTCTGTGATCTGACACACGCTCCTATCCCATCAAGAGACATGGACA 3261
DB |||||||
QY 941 AlaTyArgThrIleSerGlyThrGluThrLeuLeuSerHisGlnGluThrTrpThr 960
DB |||||||
QY 3262 ACAGATGCTCTTCATTTAGCAACATGGAGTTGTGTTAGAGATCTATGTATGCTTCT 3321
DB |||||||
QY 961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyAlaSer 980
DB |||||||
QY 3322 CTAAACAGTAATATAGAGTATATGGCATGGAGATATGATATCGATGCTTCTCGA 3381
DB |||||||
QY 981 LeuThrSerAsnIleGluValTyGlyHisGlyArgTyGlyTyArgAspAlaSerArg 1000
DB |||||||
QY 3382 GGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
DB |||||||
QY 1001 GlyTyGlyLeuSerAlaGlySerArgValArgPhe 1012
DB |||||||

```

RESULT 2

```

US-10-766-711-2
; Sequence 2, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; TITLE OF INVENTION: Pace, John
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1012

```

```

; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-2

Alignment Scores:
Pred. No.: 0 Length: 1012
Score: 5267.00 Matches: 1012
Percent Similarity: 100.0% Conservative: 0
Beat Local Similarity: 100.0% Mismatches: 0
Query Match: 66.8% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-766-711-2 (1-1012)

QY 382 ATGCAAAAGCTCTTTCCATAAGTTCTTTCTTTCAATGATCTAGCTTATTCTTCTGCTCT 441
DB |||||||
QY 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTySerCysSer 20
DB |||||||
QY 442 TTAAATGGGGGGGATATGCAGCAGAAATCATGGTTCCTCAAGGAATTTACGATGGGAG 501
DB |||||||
QY 21 LeuAsnGlyGlyGlyTyAlaAlaGluIleMetValProGlnGlyIleTyAspGlyGlu 40
DB |||||||
QY 502 ACGTTAACTGTATCATTTCCCTATCTACTGTATAGGAGATCCGAGTCGGACTACTGTTT 561
DB |||||||
QY 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
DB |||||||
QY 562 TCTGCAGGAGAGTTAACTTAAATAATCTTGACAAATCTTATTCAGCTTTGCCCTTTAAGT 621
DB |||||||
QY 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
DB |||||||
QY 622 TGTGTTTGGGAACCTTATTAGGGAGTTTACTGTTTATAGGGAGGAGCACCTCGTTGACTTTC 681
DB |||||||
QY 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
DB |||||||
QY 682 GAGAACATAGGACTTCTACAATGGGGGAGCTCTAAGTAATAGCCCTGCTGATGACTG 741
DB |||||||
QY 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120
DB |||||||
QY 742 TTTTACTATTGAGGGTTTAAAGAAATATCTTTCCAAATTTGCAATTCATTACTTCCGTA 801
DB |||||||
QY 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnSerLeuLeuAlaVal 140
DB |||||||
QY 802 CTGCTCTGTCGACCACTAATAAGGCTAGCCAGACTCCGACCTCCAGCAACAACATCTACACGCTCT 861
DB |||||||
QY 141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer 160
DB |||||||
QY 862 AATGGTACTATTATTCTAAAACAGACTCTTTGTTTACTCAATAATGAGAAGTTCTCATTC 921
DB |||||||
QY 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
DB |||||||
QY 922 TATAGTAATTTAGTCTCTCGAGATGGGGGAGCTATAGATGCTAAGAGCTTAACGCTTCAA 981
DB |||||||
QY 181 TySerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
DB |||||||
QY 982 GGAATTAGCAAGCTTTGTGTTCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
DB |||||||
QY 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
DB |||||||
QY 1042 CAAAGTAGTCACCAAGTTTCTCTGCTATGGCTAAACGAGCTCCTATGCTTTGTAGCGAAT 1101
DB |||||||
QY 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
DB |||||||
QY 1102 GTTTCAGGAGTAAGAGGGGAGGAGTGTGCTGTTTCAGATGGGACGAGGAGTGTCA 1161
DB |||||||
QY 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
DB |||||||
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTAAAGTTTTCAGAAATACTCGGCTAGAGTTT 1221
DB |||||||
QY 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
DB |||||||
QY 1222 GATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTCTCT 1281
DB |||||||
QY 281 AspGlyAsnValAlaArgValGlyGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300
DB |||||||

```

Qy	1282	AATAATCGAAAAACCTTTGTTTCTCAACAATGTTGCTTCTCTCTGTTTACATTGCTGCTAAG	1341
Db	301	AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaLys	320
Qy	1342	CAACCAACAGTGGACAGGCTTCTAATACAGTAATAATTACGGAGATCGAGAGCTATC	1401
Db	321	GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIle	340
Qy	1402	TTCTGTAAAGATGTGCGCAAGCAGGATCAAATACTCTCGATCAGTTTCTCTTGATGGA	1461
Db	341	PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly	360
Qy	1462	GAGGAGTAGTTTCTTTAGTAGCAATGTAGTCTGCGAAGAGGGAGACTATTTATGCC	1521
Db	361	GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyrAla	380
Qy	1522	AAAAAGCTCTGGTTGCTAACTGCGGCCCTGTACAAATTTTAAGGATATCGCTAATGAT	1581
Db	381	LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp	400
Qy	1582	GGTCGAGCGATTATTTAGAGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGCGAGAT	1641
Db	401	GlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAsp	420
Qy	1642	ATTATTTTCGATGGAAATCTTTAAAGAACAGCCAAAGAGAATGTCGCCGATGTTAATGGC	1701
Db	421	IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly	440
Qy	1702	GTAATCTGTCTCACAGCCATTTCCATGGGATCGGGAGGAAATAACGACATTAAAGA	1761
Db	441	ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg	460
Qy	1762	GCTAAAGCAGGCGATCAGATTCTCTTAATGATCCCATCGAGATGGCAACGGAAATAAC	1821
Db	461	AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn	480
Qy	1822	CAGCCAGCGCAGTCTTCCAAACTTCTAAAATTAACGATGGTGAAGGATACACAGGGGAT	1881
Db	481	GlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGlyTyrThrGlyAsp	500
Qy	1882	ATTGTTTTGCTAATGGAGCAGTACTTTGTACAAATCTTACGAATCTTACGATAGCAGGAGG	1941
Db	501	IleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIleGluGlnGlyArg	520
Qy	1942	ATTGTTCTTCGTGAAAGCGCAAAATTAATCAGTGAATCTCTTAAGTCAGACAGTGGGAGT	2001
Db	521	IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySer	540
Qy	2002	CTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACACACACAGCCCT	2061
Db	541	LeuTyrMetGluAlaGlySerThrTrpAspPheValThrProGlnProGlnGlnPro	560
Qy	2062	CCTCGCGCTAATCAGTTGATCAGCGCTTCCAACTCTGCAATTTGTCTCTTCTTCTTGTTA	2121
Db	561	ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeu	580
Qy	2122	GCAACCAATGCAAGTTAGAAATCTCTCCTACCAATCTCCAGCGCAAGATTTCTCATCTCGCA	2181
Db	581	AlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisProAla	600
Qy	2182	GTCATTGGTAGCACAACTGCTGGTTCTGTTACAAATTAGTGGCCCTATCTTTTTTGAGGAT	2241
Db	601	ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp	620
Qy	2242	TTGATATGATACAGCTTATGATAGTATGATTGGCTAGGCTCTAATCAAAAAATCAATGTC	2301
Db	621	LeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsnVal	640
Qy	2302	CTGAAATTTACAGTTAGGCACTAAGCCCCCAGCTAATGCCCCCATCAGATTTGACTTAGG	2361
Db	641	LeuLysLeuGlnLeuGlyThrLysProAlaAsnAlaProSerAspLeuThrLeuGly	660

```

US-10-931-779-2
; Sequence 2, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; PRIOR FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-2

Alignment Scores:
Pred. No.: 0 Length: 1012
Score: 5267.00 Matches: 1012
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 66.8% Indels: 0
DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-2 (1-1012)
QY 382 ATCCAAACGCTTCCATAGTTCTTCTTCAATGATCTTACGTTATCTTCTGCTCT 441
DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaIyrSerCysSer 20
QY 442 TTTAAATGGGGGGATATGACAGCAAAATCATGCTTCTCAAGCAATTTACGATGGGAG 501
DB 21 LeuAsnGlyGlyGlyTyraAlaAlaGluIleMetValProGlnGlyIleTyraAspGlyGlu 40
QY 502 ACGTAACTATATCCCTATCTACTGTTATAGGAGATCCGATGGGACTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGATTAAATTAATAATCTTGACAAATCTATTCGAGCTTTGCCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTCCTGGAACTTATTAGGAGTTTACTGTTTATAGGAGAGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGAGCTTCTACAAATGGGGCAGCTCTAAGTAATAGCGTGTGATGGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120
QY 742 TTTACTATTGAGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTACTTGGCGTA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCCTCTGCAACGACTTAATAAGGGTAGCCAGACTCCGACGACAAATCTACACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnLysGlySerGlnThrProThrThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTTCTAAACAGATCTTTTGTGTACTCAATAATAGAAAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyThrSerLysThrAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTATAGAGCTTAACGGTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTCTCTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220

```

```

QY 1042 CAAGTAGTCACCAGTTTCTCTGCTATGGCTAACGAGGCTCTTATTCCTTTCTAGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTCAGGAGTTAAGAGGGAGGATGCTGCTGCTTCCAGATGGGACAGGAGGTGTC 1161
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAACTTTTCCAGAAATACTCCGGTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGCACTAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACTGCTTCTTCCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTyThrSerTyThrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAAAAACCTGTTTCTCAACAATGTTGCTTCTCTGTTTACATTCGCTGAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyThrIleAlaAlaLys 320
QY 1342 CAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATTACGGAGATCGAGGAGTATC 1401
DB 321 GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyArgLysAspGlyAlaIle 340
QY 1402 TTCTGTAAAGAAATGGTGGCAAGCAGGATCCAAATAACTCTGGATCAGTTTCTTGTATGA 1461
DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly 360
QY 1462 GAGGGAGTAGTTTCTTTTAGTAGCAATGATGCTGCGGAAAGGGAGGAGTATTTATGCC 1521
DB 361 GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyraAla 380
QY 1522 AAAAGCTCTCGTTGCTTAACCTGTCGCTGTACAAATTTTAAAGGATATCGCTAATGAT 1581
DB 381 LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp 400
QY 1582 GGTGGAGCAGTATTATTAGGAGATCTCGAGAGCTCAGTTTATCTGCTGATTAAGAGAT 1641
DB 401 GlyGlyAlaIleTyThrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyThrGlyAsp 420
QY 1642 ATTATTTCGATGGGAATCTTAAAGAACACAGCCAAAGAGAGATGCTGCCGATGTTAATGCC 1701
DB 421 IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly 440
QY 1702 GTAACCTGTCTCACAAGCCATTTGATGGGATCGGGGAGGAAATTAACGACATTAAGA 1761
DB 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460
QY 1762 GCTAAAGCAGGGCATCAGATCTCTTTTAATGATCCCATCGAGATGGCAACGGAATAAAC 1821
DB 461 AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn 480
QY 1822 CAGCCAGCGCAGTCTTCCAAACTTCTAAATAATTAACGATGGTGAAGGATACACAGGGGAT 1881
DB 481 GlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGlyTyThrGlyAsp 500
QY 1882 ATTGTTTTCCTAATGGAACAGTACTTTGTGTACCAAAATGTTAGGATAGACGAGGAGG 1941
DB 501 IleValPheAlaAsnGlySerThrLeuTyThrGlnAsnValThrIleGluGlnGlyArg 520
QY 1942 ATTGTTTTCCTGAAAGGCAAAATTAATCAGTGAATTTCTTAAGTACAGACAGGTTGGAGT 2001
DB 521 IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySer 540
QY 2002 CTGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTCCACCAACCAACCAACAGCCT 2061
DB 541 LeuTyMetGluAlaGlySerThrTyAspPheValThrProGlnProProGlnGlnPro 560
QY 2062 CTTGCCCTAATCAGTTGATCAGCTTTCCTCAATCTGCAATTCGATTTCTCTTCTTCTTTA 2121
DB 561 ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeu 580

```



```
QY 2122 GCAACAAATGCGATTACGAATCTCTACCAATCTCTCAGCGCAAGATTTCTCATCTCGCA 2181
Db 581 AlaAenAenAlaValThrAenProProThrAenProProAlaGlnAenSerHisProAla 600
QY 2182 GTCATTGGTAGCACAACTGCTGGTCTGTACAAATTAGTGGCCCTATCTTTTTCAGGAT 2241
Db 601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAap 620
QY 2242 TTGGATGATACAGCTATGATAGGTATAGTATGGCTAGGTTCTAATCAAAAATCAATGTC 2301
Db 621 LeuAapAapThrAlaTyraAapArgTyrAapTyrLeuGlySerAenGlnIleAenVal 640
QY 2302 CTGAATTTACAGTTAGGACCTAAGCCCGCAGCTAATGCCCATCAGATTTGACTCTAGGG 2361
Db 641 LeuIleGlnLeuGlyThrLysProProAlaAenAlaProSerAapLeuThrLeuGly 660
QY 2362 AATGAGATGCTTAAGTATGGCTATCAAGGAAGCTCGAAGCTTGGCTGGGATCCTAATACA 2421
Db 661 AenGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAapProAenThr 680
QY 2422 GCAATAATAGTCTTACTACTCTGAAGCTACATGAGCTAAAGCTGGGTATATCTCTGGG 2481
Db 681 AlaAenAenGlyProTyrThrLeuLysAlaThrThrLysThrGlyTyrAenProGly 700
QY 2482 CTTGAGCGAGTAGCTTCTTGGTTCCTAATAGTTTATGGGATCCATTTAGATATACGA 2541
Db 701 ProGluArgValAlaSerLeuValProAenSerLeuTyrGlySerIleLeuAapIleArg 720
QY 2542 TCTGCGCATTCAGCAATCAAGCAAGTGGGATGGCGCTCTTATGTCGAGGATTATGG 2601
Db 721 SerAlaHisSerAlaIleGlnAlaSerValAapGlyArgSerTyrCysArgGlyLeuTyr 740
QY 2602 GTTCTGAGATTTCGAATTTCTTATCATGACCGCGATGCTTTTAGTCAGGATATCGG 2661
Db 741 ValSerGlyValSerAenPheTyrHisAapArgAapAlaLeuGlyGlnGlyTyrArg 760
QY 2662 TATATTAGTGGGGTATTCTTAGAGCAAACTCTACTTTGGATCATCGATGTTGGT 2721
Db 761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPheGly 780
QY 2722 CTAGCATTTACCGAAGTATTTGGTAGATCTAAGATTTATGATGTGTGTTCCAAATCAT 2781
Db 781 LeuAlaPheThrGluValPheGlyArgSerLysAapTyrValValCysArgSerAenHis 800
QY 2782 CATGCTTGATAGATCCGTTTATCTATCTACCCAAAGCTTTATGCGATTCCTATTGG 2841
Db 801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeu 820
QY 2842 TTTCGAGATGCGTTTATCCGTCGCTAGCTACGGGTTTGGGAATCAGCATATGAACCTCA 2901
Db 821 PheGlyAapAlaPheIleArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThrSer 840
QY 2902 TATACATTTGCAGAGAGAGCGATGTTCTGTTGGGATAATAACTGTCTGCTCGAGAGATT 2961
Db 841 TyrThrPheAlaGluGluSerAapValArgTyrAapAenAenCysLeuAlaGlyGluIle 860
QY 2962 GGAGCGGGATTACCGATTGTGATTCTCATCTAAGCTCTATTGAATGAGTTGCGTCT 3021
Db 861 GlyAlaGlyLeuProIleThrProSerLysLeuTyrLeuAenGluLeuArgPro 880
QY 3022 TTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAAGCGATCAA 3081
Db 881 PheValGlnAlaGluPheSerTyrAlaAapHisGluSerPheThrGluGluGlyAapGln 900
QY 3082 GCTCGGGCATCAAGAGCGGACATCTCCTAATCTATCATAGTTCTCTGTTGAGTGAAGTTT 3141
Db 901 AlaArgAlaPheLysSerGlyHisLeuLeuAenLeuSerValProValGlyValLysPhe 920
QY 3142 GATCGATGTTCTAGTACACATCTTAATAAATATAGCTTTATGGCGGCTTATATCTGTGAT 3201
Db 921 AapArgCysSerSerThrHisProAenLysTyrSerPheMetAlaAlaTyrIleCysAap 940
QY 3202 GCTTATCGCACCATCTCTGGTACTGAGACACGCTCTCTATCCATCATCAAGAGACATGGACA 3261
```

```
Db 941 AlaTyrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTyrThr 960
QY 3262 ACAGATGCTTTTCATTATTAGCAAGACATGGAGTTGTGGTTAGAGGATCTATGTATGCTTCT 3321
Db 961 ThrAapAlaPheHisLeuAlaAargHisGlyValValValArgGlySerMetTyrAlaSer 980
QY 3322 CTAAACAGTATATAGAGTATATGGCCATCGAAGATATAGTATCGAGATGCTTCTCGA 3381
Db 981 LeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAapAlaSerArg 1000
QY 3382 GGCATGCTTTTCAGTGCAGGAAGTAGATCCGGTTC 3417
Db 1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012

RESULT 4
US-10-467-534-9
; Sequence 9, Application US/10467534
; Publication No. US20040131625A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-9

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
D8: 4 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-467-534-9 (1-1013)
QY 382 ATGCAACGTCCTTTCCATAAGTCTTCTTCTTTCAATGATCTAGCTTATCTTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
QY 442 TTAATATGGGGGGGATATGCAGCAAAATCATGGTCTCTCAAGGAATTTACCATGGGAG 501
Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGingIlyIleTyrAapGlyGlu 40
QY 502 ACGTTAACTGTATCAATTTCCCTATACTGTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAapProSerGlyThrThrValPhe 60
QY 562 TCTCGAGAGAGTTAACTTAAAAATCTTGACAAATCTTATGCAAGCTTTGCGCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAenLeuAapAenSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTGTTTGGGAACCTTATTAGGAGTTTTACTGTTTTAGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACATG 741
Db 101 GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAapSerAlaAenSerGlyLeu 120
```

	Qy	742	T T T A C T A T T G A G G G T T T T A A A G A A T T A T C C T T T T C C A A T T G C A A T T C A T T A C T T G C C G T A	801
	Dd	121	P h e T h r I l e G l u G l y P h e L y s G l u L e u S e r P h e S e r A s n C y b A n S e r L e u L e u A l a V a l	140
	Qy	802	C T G C C T G C T G C A A C G A C T A A T A A G G G T A G C A G A C T C C G A G A C A A C A T C T A C A C G T C T	861
	Dd	141	L e u P r o A l a A l a T h r T h r A s n A s n G l y S e r G l n T h r P r o T h r T h r S e r T h r P r o S e r	160
	Qy	862	A A T G G T A C T A T T A T T A C T A A A C A G A T C T T T G T T A C T C A A T A A T A G A A G T T C T C A T T C	921
	Dd	161	A s n G l y T h r I l e Y r S e r L y s T h r A s p L e u L e u L e u L e u A s n G l u L y s P h e S e r P h e	180
	Qy	922	T A T A G T A A T T T A G T C T C T G G A G A T G G G G A G C T A T A G A T C T A A G A G C T T A A C G G T T C A A	981
	Dd	181	T y r S e r A s n L e u V a l S e r G l y A s p G l y A l a I l e A s p A l a L y s S e r L e u T h r V a l G i n	200
	Qy	982	G G A A T T A G C A G C T T T G T C T T C C A G A A A A T A C T G C T C A A G C T A N T G G G G A G C T T G T	1041
	Dd	201	G l y I l e S e r L y s L e u C y s V a l P h e G l n G l u A s n T h r A l a G l n A l a A s p G l y G l y A l a C y s	220
	Qy	1042	C A A G T A G T C A C A G T T T C T C T G C T A T G G C T A A C G A G G C T C T A T T G C C T T T G T A G C G A A T	1101
	Dd	221	G i n V a l V a l T h r S e r P h e S e r A l a M e t A l a A s n G l u A l a P r o I l e A l a P h e I l e A l a A s n	240
	Qy	1102	G T T C C A G G A G T A A G A G G G G A G G A T T G C T G C T T T C A G A T G G C A G C A G G A G A G T G T C A	1161
	Dd	241	V a l A l a G l y V a l A - g G l y G l y I l e A l a A l a V a l G l n A s p G l y G l n G l y V a l S e r	260
	Qy	1162	T C A T C T A C T T C A A C A G A G A T C C A G T A G T A A G T T T T C C A G A A T A C T C G G T A G A G T T T	1221
	Dd	261	S e r S e r T h r S e r T h r G l u A s p P r o V a l V a l S e r P h e S e r A r g A n T h r A l a V a l G l u P h e	280
	Qy	1222	G A T G G G A A C G T A G C C G A G T A G G A G G A G G A T T T A C T C C T A C G G E A A C G T T G C T T C C T G	1281
	Dd	281	A s p G l y A s n V a l A - a a r g v a l G l y G l y I l e Y r S e r T y r G l y A s n V a l A l a P h e L e u	300
	Qy	1282	A A T A A T G G A A A A C C T T G T T T C C A A C A A T G T T C C T C C T G T T T A C A T T G C T G C T A A G	1341
	Dd	301	A s n A e n G l y L y s T h r L e u P h e L e u A s n A s n V a l A l a S e r P r o V a l T y r I l e A l a A l a G l u	320
	Qy	1342	C A A C C A A C A C T G G A C G C T T C T A A T A C A G A T A A T A A T A C G C A G A T G A G G A G A G C T A T C	1401
	Dd	321	G l n P r o T h r A e n G l y G l n A l a S e r A s n T h r S e r A s p A n T y r G l y A s p G l y G l y A l a I l e	340
	Qy	1402	T T C T G T A A G A A T G G T G C G C A A - - - C G A G A T C C A A T A A C T C T G G A T C A G T T T C C T T T G A T	1458
	Dd	341	P h e C y s L y s A e n G l y A l a G l n A l a A l a G l y S e r A e n A e n S e r G l y S e r V a l S e r P h e A s p	360
	Qy	1459	G G A G A G G A G T A G T T T C T T T A G T A G A A T C T A G C T G C T G G A A G G G G A G C T A T T A T	1518
	Dd	361	G l y G l u G l y V a l V a l P h e P h e S e r A s n V a l A l a A l a G l y L y s G l y G l y A l a I l e Y r	380
	Qy	1519	G C C A A A A A G C T C C G G T T G C T A A C T T G C C C C T G T A C A A T T T T T A A G A A A T A T C G C T A A T	1578
	Dd	381	A l a L y s L y s L e u S e r V a l A a s n C y s G l y P r o V a l G l n P h e L e u G l y A s n I l e A l a A s n	400
	Qy	1579	G A T G T G G A G C A T T A T T T T A G G A A T C T G G A G A G C T C A G T T A T T C T G C T G A T T A T T G A	1638
	Dd	401	A s p G l y G l y A l a I l e Y r L e u G l y G l u S e r G l y G l u L e u S e r L e u S e r A l a A s p T y r G l y	420
	Qy	1639	G A T A T T A T T T C A T G G G A A T C T T A A A A G A C A C C A A G A G A A T G C T C C G A T T A A T	1698
	Dd	421	A s p I l e I l e P h e A s p G l y A s n L e u L y s a r g T h r A l a L y s G l u A s n A l a A l a A s p V a l A s n	440
	Qy	1699	G C G C T A A C T G C T C C T C A C A A G C C A T T T C A T G G A T C G G A G G G A A A T A A C A C A T T A	1758
	Dd	441	G l y V a l T h r V a l S e r S e r G l n A l a I l e S e r M e t G l y S e r G l y G l y L y s I l e T h r T h r L e u	460
	Qy	1759	A G A C T T A A G C A G G G C A T C A G A T T C T T T A T G A T C C A T C G A T G C G A A C G G A A T	1818
	Dd	461	A r q A l a L y s A l a g l y H i s G l n I l e L e u P h e A s n A s p P r o I l e G l u M e t A l a A s n G l y A s n	480

1819	QY	1819	AACAGCCAGCGAGTCTTCTTCAAACTCTTAAATAATTAACGATCGTGAAGATACACAGGG	1879
	DB	481	AsnGlnProAlaGlnSerSerGluProLeuLeuIleAsnAspGlyGluGlyTyrThrGly	500
	QY	1879	GATATTCCTTTTGTCTAATGAAGCAGTACTTGTGTACCAAAATGTTACGATAGACAGCAAGGA	1938
	DB	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly	520
	QY	1939	AGGATTCCTTCTGTGAAAAAGGCAAAATTATCAGTCGAATTTCTCTAAAGTCAGACAGGTGGG	1998
	DB	521	ArgIleValLeuArgGluLeuAlaLeuLeuSerValAsnSerLeuSerGlnThrGlyGly	540
	QY	1999	AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACAAACCCACACACACAG	2058
	DB	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln	560
	QY	2059	CCTCTCGCGCTAATCAGTTGTATCAGTCACAGCTTTTCCAATCTGCATTTGTCTCTTTCTTTTG	2118
	DB	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu	580
	QY	2119	TTAGCAAAACAATGCAGTTTACGAATCTCTCTCAACCAATCTCTCAGCGCAAGATTCTCATCCT	2178
	DB	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
	QY	2179	CGAGTCATGTTAGCACCAACTGCTGGTCTGTCTTCAATATTAGTGGCGCTATCTTTTTCAG	2238
	DB	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
	QY	2239	GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTCTTAATCAAAAAATCAAT	2298
	DB	621	AspLeuAspAspThrAlaIleTyrAspArgTyrAspTyrLeuGlySerAsnGlnLeuIleAsp	640
	QY	2299	GTCTCTGAAATATACAGTTTAGGAGCTAAAGCCCCAGCTAATGCCCATCAGATTTCAGCTCTA	2358
	DB	641	ValLeuLeuLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
	QY	2359	GGGAATGAGATGCCTAAGTATGGCTATCAAGGAAGCTGGGAAGCTTGGTGGGATCCCTAAT	2418
	DB	661	GlyAsnGluMetProLeuTyrGlyTyrGlnGlySerTyrTrpLeuAlaIleTrpAspProAsn	680
	QY	2419	ACAGCAAAATAATGTCCTTATCTGAAAGCTACATGAGCTAATAAAGCTGGGTATAATCCT	2478
	DB	681	ThrAlaAsnAsnGlyProTyrThrLeuLeuAlaThrTrpThrLeuTyrGlyTyrAsnPro	700
	QY	2479	GGGCTCGAGCAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTTATAGATATA	2538
	DB	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle	720
	QY	2539	CGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTTCGAGGATTA	2598
	DB	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740
	QY	2599	TGGGTCTTCGGAGTTTCGAAATTTCTTCTCATCATGACCGCATGCTTTAGGTTCAGGATAT	2658
	DB	741	TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760
	QY	2659	CGGTATATTAGTGGGGGTATTCTCTTAGGAGCAACTCTCTACTTTGGATCATTCGATGTTT	2718
	DB	761	ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780
	QY	2719	GGTCTAGCAATTACCGAAGTATTTGGTATAGCTAAGAATATATGATGTGTCTGTTCCCAAT	2778
	DB	781	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn	800
	QY	2779	CATCATGCTTGATAGGATCGGTTTATCTATCTACCAACAGCTTTATGTGGATCCCTAT	2838
	DB	801	HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr	820
	QY	2839	TTGTTTCGGAGATGCGTTTATCCGTCCTAGCTACGGGTTTGGGAATCAGCATATGAAACC	2898
	DB	821	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840
	QY	2899	TCATATACATTTGCAGAGGAGAGCGAATGTTCTGTTGGGATATAATATCTCTCGCTGGAGAG	2958

```
Db      841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAenAenCysLeuValGlyGlu 860
Qy      2959 ATTGAGCGGGAATTACCGATTGTGATTACTCCATCTAAGCTCTATTCAATGAGTTCGGT 3018
Db      861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
Qy      3019 CCTTTCGTGCAAGCTGAGTTTCTTATGCGGATCATGAAATCTTTTACAGAGAGGCCAAT 3078
Db      881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy      3079 CAAGCTCGGGCATTCACAGCGGCATCTCCTAAATCTCATGATCCGTTCCGTGGAGTGAAG 3138
Db      901 GlnAlaArgAlaPheAenSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
Qy      3139 TTTCATCATGTTCTTATGACACATCCTTAATAATATAGCTTTATGGCGCTTATATCTGT 3198
Db      921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy      3199 GATGCTTATCGCACCATCTCTGGTACTGAGACAAACGCTCCTATCCATCAAGACATGG 3258
Db      941 AspAlaTyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960
Qy      3259 ACAACAGATCGCTTTCATTACCAAGACATCGAGTTGTGTAGAGGATCTATGATGCT 3318
Db      961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgLysMetTyrAla 980
Qy      3319 TCTCTAACAAAGTAATATAGAAGTATATGCCCATGAAGATATGATATCGAGATCTTCT 3378
Db      981 SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy      3379 CGAGCTATGGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db      1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

RESULT 5

```
US-10-498-327-5
; Sequence 5, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-5
```

Alignment Scores:

```
Pred. No.: 0 Length: 1013
Score: 5131.50 Matches: 983
Percent Similarity: 98.8% Conservative: 18
Best Local Similarity: 97.0% Mismatches: 11
Query Match: 65.1% Indels: 1
DB: 5 Gaps: 1
```

US-10-701-844-1 (1-4435) x US-10-498-327-5 (1-1013)

```
Qy      382 ATGCAAAAGCTCTTCCATGAGTTCTTCTTCAATGATCTAGCTTATCTTCGTCTCT 441
Db      1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
Qy      442 TTAATATGGGGGGGATATCAGACAGAAATCATGTTCTCTCAAGGAATTTACCATGGGAG 501
Db      21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
```

```
Qy      502 ACGTTAACTGTATCATTTCCCTATACGTGTATAGGAGATCCGAGTGGAGCTACTGTTTTT 561
Db      41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy      562 TCTCAGAGAGATTAAACATTAAAAAATCTTGACAAATCTATTGACGCTTTCCTCTTAAGT 621
Db      61 SerAlaGlyGluLeuThrLeuLysAenLeuAenSerIleAlaAlaLeuProLeuSer 80
Qy      622 TGTTTTGGAACTTATTAGGAGGTTTTTACTGTGTTTTAGGAGAGGACACTCGTTGACTTTC 681
Db      81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy      682 GAGAACATACGAGACTTCTCAATATGGGCGAGCTCTAAGTAATAGCGCTGTGTAGGAGCTG 741
Db      101 GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
Qy      742 TTTTACTATGAGGCTTTTAAAGAAATATCCTTTTCCAAATTCGAATTCATTTACTTGCCTA 801
Db      121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
Qy      802 CTGCTCTGTCACACGACTAATAAGGGTAGCCAGACTCCGACGACACAACTCTACACCGTCT 861
Db      141 LeuProAlaAlaThrThrAenAenGlySerGlnThrProThrThrSerThrProSer 160
Qy      862 AATGGTACTATTATTCTTAAACAGACTTTTGTGTTACTCAATATAGAGAGTTCTCATTC 921
Db      161 AenGlyThrIleTyrSerLysThrAspLeuLeuLeuAenAenGlnLysPheSerPhe 180
Qy      922 TATAGTATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA 981
Db      181 TyrSerAenLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
Qy      982 GGAATATTAGCAAGCTTTGTGCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
Db      201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy      1042 CAAGTAGTCACAGTTTCTGCTATGCTTAACGAGGCTCTATTCCTTGTAGCGAAT 1101
Db      221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240
Qy      1102 GTTCAGGAGTAAGAGGGGAGGATGCTCTGTTCAAGATGGGACGACGAGGTGTCA 1161
Db      241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAenGlyGlnGlnGlyValSer 260
Qy      1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTTCAGAAATACTCGCGGTAGAGTTT 1221
Db      261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
Qy      1222 GATGGAAACGTAGCCCGAGTAGGAGGAGGATTTACTCTTACGGGAACGTTGCTTCTCTG 1281
Db      281 AspGlyAenValAlaArgValGlyGlyIleTyrSerTyrGlyAenValAlaPheLeu 300
Qy      1282 AATAATGGAAAAACCTGTTTCTCAACAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1341
Db      301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaGlu 320
Qy      1342 CAACCAACAAGTGGACAGGCTTCTAATACGAGTAATAATACGAGAGATGGAGGAGCTATC 1401
Db      321 GlnProThrAenGlyGlnAlaSerAenThrSerAenThrSerAenThrGlyAspGlyAlaIle 340
Qy      1402 TTCTGTAGAATGGTGGCGCAA---GCAGGATCCAAATAACTCTGGATCAGTTTCCTTGAAT 1458
Db      341 PheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAsp 360
Qy      1459 GGAGAGGAGTAGTTTTCTTTTAGTAGCAATGTAGCTGTGGGAAAGGGAGCTATTTAT 1518
Db      361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTyr 380
Qy      1519 GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
Db      381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
```

1579 GATGGTGGAGCGATTATTTAGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
1580 |||||
401 AspGlyGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420
1639 GATATATTTTCGATGGGAATCTTAAAGAAACAGCCAAAGAGAAATGCTGCCGATGTTAAT 1698
1640 |||||
421 AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
1699 GCGTAACTGCTCTCACAAGCCATTTTCAGTGGGATCGGAGGGGAATAACGACATTA 1758
1700 |||||
441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrLeu 460
1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGGAAT 1818
1819 |||||
461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
1819 AACACGCCAGCGAGCTCTTCCAAACTTCTAAATTAACGATGGTGAAGATACACAGGG 1878
1880 |||||
481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyrThrGly 500
1879 GATATTTGTTTTCGTAATCGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCA 1938
1939 |||||
501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
1939 AGGATTGTTCTTCGTAAGGCAAAATATCAGTCAATCTCTAAGTCAGACAGAGTGG 1998
1999 |||||
521 ArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
1999 AGTCTCTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCACCAACACACAG 2058
2059 |||||
541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
2059 CTTCTGCGCCTTAATCAGTTGATCAGCTTTCAATCTGCAATTTGCTCTTCTCTTTG 2118
2119 |||||
561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
2119 TTAGCAAAACAATCAGTTACGATCTCTTACCAATCTCTACCAAGCAAGATCTCATCT 2178
2179 |||||
581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
2179 GCATCATTTGTTAGCACAACTGCTGGTCTGTTTCAATTAAGTGGGCTATCTTTTGTAG 2238
2239 |||||
601 AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
2239 GATTGGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2298
2299 |||||
621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
2299 GTCTGAAATTCAGTTAGGAGCTTAAGCCCCCAGCTAAATGCCCATCAGATTTGACTCTA 2358
2359 |||||
641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
2359 GCGAATCAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCCTTAAT 2418
2419 |||||
661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
2419 ACAGCAAAATATGCTCTTACTCTGAAAGCTACATGAGCTAAACTGGGTATATCCT 2478
2479 |||||
681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
2479 GGGCTCAGCAGTATGCTCTTTGTTTCCAAATGTTTATGGGATCCATTTAGATATA 2538
2539 |||||
701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
2539 CGATCTCGCATTCAGCAATTCAGCAAGTGTGATGGGCTCTTATTTGTCAGGATTA 2598
2599 |||||
721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
2599 TGGGTTTTCGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTAGTTCAGGATAT 2658
2659 |||||
741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
2659 CGGTATATAGTGGGGTTATCTCTAGGACCAACTCTCTTGTGATCATCATGATGTTT 2718

761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCCAAT 2778
2779 |||||
781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTTATGTGATCCCTAT 2838
2839 |||||
801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
2839 TTGTTCCGAGATGCGTTTATCCGCTAGCTACGCGTTTGGGAATCAGCATATGAACCC 2898
821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
2899 TCATATACATTTGCAGAGAGAGCGATGTTCTCTGGGATAATAACTGTCTGGCTGGAGAG 2958
841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
2959 ATTGAGCGGGATTTACCGATTGCTGATTCTCCATCTAAAGCTCTATTTGAATGAGTTCGT 3018
861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
3019 CTTCTCGTCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGAT 3078
881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGlyAsp 900
3079 CAAGCTCGGCAATTCAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTGGAGTGAAG 3138
901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
3139 TTGTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTTATGGCGCTTATATCTGT 3198
921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
3199 GATGCTTATCGCACCATCTCTGCTACTGAGACACCGCTCTCTATCCCATCAAGAGCATGG 3258
941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
3259 ACAACAGATGCTTTCATTTAGCACACATGAGTGTGTGTAGAGCATCTATGATGCT 3318
961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
3319 TCTCTAACCAAGTAATATAGAGTATATGCGCATGAGGATATGAGTATCGAGATGCTTCT 3378
981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
3379 CGAGGCTATGGTTTGTAGTGCAGGAAGTAGAGTCCGGTTC 3417
1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 6

US-10-498-327-81
; Sequence 81, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 1013
; TYPE: PRF
; ORGANISM: Chlamydia trachomatis
US-10-498-327-81
Alignment Scores:

Pred. No.:	0	Length:	1013
Score:	5131.50	Matches:	983
Percent Similarity:	98.8%	Conservative:	18
Best Local Similarity:	97.0%	Mismatches:	11
Query Match:	65.1%	Indels:	1
DB:	5	Gaps:	1

US-10-701-844-1 (1-4435) x US-10-498-327-81 (1-1013)

QY	382	ATGCAAGCTCTTCCATAAGTCTTCTTCAATGATTTCTAGCTTATTTCTGCTCT	441
DB	1	MetGlnThrSerPheHisLeuPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer	20
QY	442	TAAATGGGGGGATATGCACAGAAATCATGGTCTCTCAAGGATTTACATGGGGAG	501
DB	21	LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu	40
QY	502	ACGTTAACTGATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGCACTACTGTTTT	561
DB	41	ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGAGGAGGTTAAACATTTAAATAATCTTGCAATTTCTATTGCACTTTGCTTTAAGT	621
DB	61	SerAlaGlyGluLeuThrLeuLeuAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer	80
QY	622	TGTTTTGGAACTTATTAGGGAGTTTACTGTTTTAGGGAGGAGACACTCGTTGACTTTC	681
DB	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
QY	682	GAGACATACGACTTCTACAAATGGGCGAGCTTAAGTAAATAGCGCTGCTGATGACTG	741
DB	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTGAGGGTTTAAAGATTTATCTTTTCCAAATTGCAATTTACTTCCCGTA	801
DB	121	PheThrIleGluGlyPheIleGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGTGCAACGACTTAATAGGGTAGCCAGACTCCGACGACAACTCTACCGTCT	861
DB	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer	160
QY	862	ATGCTACTATTATTCTAAACAGACTTTTGTGTTACTCAATAATGAGAGTTCATTC	921
DB	161	AsnGlyThrIleTyrSerIleThrAspLeuLeuLeuAsnAsnGlnIlePheSerPhe	180
QY	922	TATAGTAATTAGTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA	981
DB	181	TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaIleAspSerLeuThrValGln	200
QY	982	GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT	1041
DB	201	GlyIleSerLeuLeuCysValPheGlnGlnAsnThrAlaGlnAlaAspGlyGlyAlaCys	220
QY	1042	CAAGTAGTCAACAGTTTCTGCTATGCTTAACGAGCTCTTATTCCTTTCTGAGCAAT	1101
DB	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	GTTTCAGGAGTAGAGGGGAGGATTGCTGCTGTTTCAGGATGGGAGGAGGTGCA	1161
DB	241	ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer	260
QY	1162	TCATCTACTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCGGTAGAGTTT	1221
DB	261	SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGAGCTAGCCGACTAGGAGGAGGATTACTCTACGGAGCTGCTTTCCCTG	1281
DB	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300
QY	1282	AATAATGGAAAAACCTTTGTTCTCAACAATGTTGCTTCTCTGTTTACATGCTGCTAAG	1341
DB	301	AsnAsnGlyIleThrLeuPheLeuAsnValAlaSerProValTyrIleAlaAlaGlu	320

QY	1342	CAACCAAGTGCACAGGCTTCTTAATACGAGTAATAATACGAGATGGAGGAGCTATC	1401
DB	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle	340
QY	1402	TTCTGTAAAGATGCTGGCCAA---GCAGGATCCATAAATCTCTGGATCAGTTTCTTGTAT	1458
DB	341	PheCysIleAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGAGTAGTTTCTTTTAGTACAAATGCTGCTGGGAAAGGGGAGCTATTAT	1518
DB	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIleGlyGlyAlaIleTyr	380
QY	1519	GCCAAAAGCTCTCGTTTCTAACTGTGGCCCTCTACAATTTTAAAGAAATATCGCTAAT	1578
DB	381	AlaIleIleLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGCTGAGCGATTTATTATTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATATCGA	1638
DB	401	AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly	420
QY	1639	GATATATTTCGATGGGAATCTTAAAGAACACGCCAAAGAAATGCTGCCGATGTTAAT	1698
DB	421	AspIleIlePheAspGlyAsnLeuLeuArgThrAlaIleGluAsnAlaAlaAspValAsn	440
QY	1699	GGCGTAACTGTCTCACAAGCCATTTTCGATGGATCGGAGCGGAGAAATTAACGACATTA	1758
DB	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu	460
QY	1759	AGAGCTAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAAAACGAAAT	1818
DB	461	ArgAlaIleAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACGAGCAGCGCGAGCTTCCAACTCTTAAATAATTAACGATGGTGAAGGATACACAGG	1878
DB	481	AsnGlnProAlaGlnSerSerGluProLeuIleAsnAspGlyGlyIleThrThrGly	500
QY	1879	GATATGTTTTCGCTTAAGGAGCAGTACTTTGTACCAAAATGTTTACCATAGACGACGA	1938
DB	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATGTTCTCTGTTGAAAAGCGCAAAATTAATCAGTGAATTTCTTAAGTCACAGGTGG	1998
DB	521	ArgIleValLeuArgGlyAlaIleLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGATATGGAAGCTGGGAGTACATGGGATTTGTAACTTCCACACACCACACAG	2058
DB	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTGCGCTTAATCAGTTGATCAGCTTTCCAACTGCTGCTTCTCTTCTCTTCTTG	2118
DB	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerSerSerLeu	580
QY	2119	TTAGCAAAATGCAAGTTACGAATCTCTTACCAATCTCCAGCCGCAAGATTTCTCATCT	2178
DB	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	GCAGTCATTGTTAGCACAACTGCTGCTGTTTCTTACAAATAGTGGGCTTCTTTTTGAG	2238
DB	601	AlaIleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTTGGATGATACAGCTTATGATAGGTATGATGCTAGGTCTTAATCAAAAATCAAT	2298
DB	621	AspLeuAspAspThrAlaTyrAspArgTyrAspThrLeuGlySerAsnGlnIleAsp	640
QY	2299	GTCTCGAAATACAGTTAGGAGTAAAGCCCGCCAGCTTAATGCCCCATCAGATTGCTCTA	2358
DB	641	ValLeuIleLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATGAGATGCCCTAAGTATGCTGCTATCAGGAAGCTGGAAGCTGCGGTGATCTAAT	2418
DB	661	GlyAsnGluMetProIleTyrGlyTyrGlnGlySerTyrLeuLeuAlaTyrAspProAsn	680
QY	2419	ACAGCAAAATATGGTCTTATCTCTGAAAGCTACATGCACTAAAACTGGGTATAATCCT	2478

681	Db	ThrAlaAsnAenGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro	700
2479	QY	GGGCGCTGAGCGAGTACGCTCTTTGGTTCCAAATAGTTTATGGGATCCAAATTTAGATATA	2538
701	Db	GlyProGlnArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle	720
2539	QY	CGATCTCGCATTCAGCAATTCAAAGCAAGTGTGGATGGGGCTCTTATTGTGCGAGATTAT	2598
721	Db	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740
2599	QY	TGGGTTCTCGAGTTTCGAATTTCTCTATCATGACCGCATCTCTTTAGTTCAGGATAT	2658
741	Db	TrpValSerGlyValSerAsnPhePheThiAspArgAspAlaLeuGlyGlnGlyTyr	760
2659	QY	CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCCTACTTTGGATCATCGATGTTT	2718
761	Db	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780
2719	QY	GGTCTACATTTACCGAAGTATTGGTAGATCTAAAGATTATCTAGTGTGCTCCCAAT	2778
781	Db	GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn	800
2779	QY	CATCATGCTTGCATAGGATCCGTTTATCTATCTACCCAAACAAGCTTTATGTGGATCCTAT	2838
801	Db	HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr	820
2839	QY	TTGTTCCGAGATGCGTTTATCCGTGTAGCTACCGGTTTGGGAATCAGCATATGAATACC	2898
821	Db	LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr	840
2899	QY	TCATATACATTTGCAGAGGAGCGATGTTGCTTGGGATAATAACTCTCTGGCTGGAGAG	2958
841	Db	SerTyrThrPheAlaGluGluSerAspValArgTrpAsnAsnLysLeuValGlyGlu	860
2959	QY	ATTCGAGCGGATTTACCGATTGTGATTATCTCCATCTAAGCTCTATTTGAATGATGTCGT	3018
861	Db	IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg	880
3019	QY	CGTTTCGTGCAAGCTGAGTTTCTTATGCCGATCATCAATCTTTTACAGAGGAGGCGAT	3078
881	Db	ProPheValGlnAlaGluPheSerTyrAlaAsnHisGluSerPheThrGluGluGlyAsp	900
3079	QY	CAAGCTCGGGCATTCAGAGCGGACATCTCCTAAATCTATCAGTTCCTGTGGAGTGAAG	3138
901	Db	GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys	920
3139	QY	TTTGATCGATGTTCTAGTACACATCTCAATAAATATAGCTTTATGGCGGCTTATCTGT	3198
921	Db	PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys	940
3199	QY	GATGCTTATCGCACCATCTCTGTACTGAGACAACGGTCTCTATCCCATCAAGACATGG	3258
941	Db	AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp	960
3259	QY	ACAAAGATGCGCTTTCATTATAGCAAGACATGGATTGTGGTTAGAGGATCTATCTATGCT	3318
961	Db	ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla	980
3319	QY	TCCTCTAACAAAGTAAATATAGAAGTATATGCCCATGGAAGATATGAGTATCGAGATGCTTCT	3378
981	Db	SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer	1000
3379	QY	CGAGCTATGTTTTCAGTGCAGGAAGTAGATCCGGTTTC	3417
1001	Db	ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe	1013

QY 1102 GTTCGAGAGTAAGAGGGGAGGAGTCTGCTGCTTTCAGGATGGCAGCAGGGAGTGTCA 1161
Db ValAlaGlyValArgGlyGlyLeuAlaGlyLeuAlaValGlnAspGlyGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTGTTCAGAAATACCTGCGGTAGAGTTT 1221
Db SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGAAACGTAGCCGAGTAGGAGGAGGATTTACTCTCGGGAACGTTGCTTTCCTG 1281
Db AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGAAAAACCTTGTTTCTCAACAATGTTGCTTCTCTGTTTACATGCTGCTAAG 1341
Db AsnAsnGlyLeuThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320
QY 1342 CAACCAACAGTGGACAGCTCTAATACGAGTAATAATACGAGAGATGGAGGAGCTATC 1401
Db GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340
QY 1402 TTCTGTAAGAAATGTCGCAAA---GCAGGATCCAATAACTCTGATCAGTTTCTTTGAT 1458
Db PheCysLeuAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGAGGAGTAGTTTCTTTAGTAGCAATGATAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
Db GlyGluGlyValAlaPhePheSerSerAsnValAlaAlaGlyLeuGlyAlaIleTyr 380
QY 1519 GCCAAAAAGCTCTCGGTGCTAACTGTCGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
Db AlaIleLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATGCTGAGCGCAATTTATTTAGAGAACTCGAGAGCTCAGTTTATCTGCTGATTAATGGA 1638
Db AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAATGCTGCCGATGTTAAT 1698
Db AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGCTAACGTGCTCTCAAGCCATTTGATGGGATCGGAGGAGGAAATAACGACATTA 1758
Db GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGATCAGATTTCTCTTAAATGATCCCATCGAGATGGCAACGGAAT 1818
Db ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACCGCCAGCCAGTCTTCCAAACCTTCAAAAAATTACGATGGTGAAGGATACACAGG 1878
Db AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
QY 1879 GATATGTTTGTCTAAAGGAGCAGTCTTGTACCAAAATGTTACGATAGCAGCAAGGA 1938
Db AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGCATGTTCTTCGTGAAAGGCAAAATATCAGTGAATTCCTAAGTCAGACAGAGTGG 1998
Db ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTCCCAACACCAACACAG 2058
Db SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTGCGCGCTAACTCAGTTGATCAGCTTTCCAATCTGCATTTGCTCTTCTCTTTG 2118
Db ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAAAACAATGCAATTCCTCAATCCTCAAAATCCTCCAGCGCAAGATTCATCCT 2178
Db LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCATTTGGTAGCAAACTGCTGTTCTGTATCAATTAGTGGGCTATCTTTTGGAG 2238

Db 601 AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTCTTAATCAAAAATCAAT 2298
Db AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTGAAATTTACAGTTAGGAGCTAAGCCCCACCTAATGCCCATCAGATTTGACTCTA 2358
Db ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGATGAGATGCTTACCTAGTATGGCTATCAAGCAAGCTGGAAGCTTGGCTGGATCCTAAT 2418
Db GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAATAATGCTCTTATCTACTGAAAGCTCATGCACTAAACCTGGGTATAATCCT 2478
Db ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCTGAGCAGTAGCTTCTTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
Db GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTGTGCGAGATTA 2598
Db ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGGATTTTCGAATTTCTTATCATGACCGCGATGCTTTTAGTCAGGATAT 2658
Db TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTTAGTGGGGTTTATCTTAGGAGCAAACTCTTACTTTTGGATTCATCGATGTTT 2718
Db ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACGAAGTATTTGGTAGACTAAAGATTAATGATGTGTCGTTCCCAAT 2778
Db GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTGCATAGGATCGGTTTATCTATCTACCCCAACAGCTTTATGCGATCCTAT 2838
Db HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTTCGAGATGCGTTTATCCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAAACC 2898
Db LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCGAGAGGAGCGATGCTTCGTTGGGATAATACTGCTGCTCGAGAG 2958
Db SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGGATTTACCGATTTGTGATTACTCCATCTAAGCTCTATTTGAATGAGTTGCGT 3018
Db IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTGCAAGCTGAGTTTCTTTATCCGATCAATGAATCTTTTACAGAGAAGCGGAT 3078
Db ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAGCTTCGGCATTTCAAGAGCGGACATCTCTTAATCTATCAGTTTCTGTTGGAGTGAAG 3138
Db GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCAGATGTTCTAGTACACATCTCTAATAATAGCTTTATCGCGCTTATCTGT 3198
Db PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATGCGCAATCTCTGGTACTGAGACAACGCTCTTATCCATCCATCAAGAGCATGG 3258
Db AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGlnThrTrp 960
QY 3259 ACAACAGATGCTTTCATTTAGCAAGACATGAGTGTGTTAGGAGTATGATGTATGCT 3318

Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValAlaGlySerMetTyrZala 980
 Qy 3319 TCCTCAACAGTATATAGAGTATATGCGCATCGAAGATATGAGTATCGAGATGCTTCT 3378
 Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 Qy 3379 CGAGCGTATGTTTTCAGTGCAGGAGTATGAGTCCGGTTC 3417
 Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerIleValArgPhe 1013

RESULT 8

US-10-498-327-9
 ; Sequence 9, Application US/10498327
 ; Publication No. US20050106162A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Grandi, Guido
 ; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
 ; FILE REFERENCE: 002441.00085
 ; CURRENT APPLICATION NUMBER: US/10/498,327
 ; CURRENT FILING DATE: 2004-06-10
 ; PRIOR APPLICATION NUMBER: PCT/IB02/05761
 ; PRIOR FILING DATE: 2002-12-12
 ; NUMBER OF SEQ ID NOS: 262
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 9
 ; LENGTH: 1013
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis
 US-10-498-327-9

Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5123.50 Matches: 979
 Percent Similarity: 98.8% Conservative: 22
 Best Local Similarity: 96.6% Mismatches: 11
 Query Match: 65.0% Indels: 1
 DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-9 (1-1013)

Qy 382 ATGCAAGCTCTTTCATAGTCTTCTTCAATGATCTAGCTTATCTTGTGCTCT 441
 Db 1 MetGlnThrSerPheHisIleValLeuSerMetIleLeuAlaTyrSerCysSer 20
 Qy 442 TTAATGGGGGGGATATCGACAGCAAAATCATGTTCTTCAAGGAATTCAGATGGGAG 501
 Db 21 LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
 Qy 502 AGCTTAATCTGATCATTTCCCTATGTTATAGGATCCGAGTGGGACTACTGTTT 561
 Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 Qy 562 TCTGACGAGAGCTTAACATTAATAAATCTTGCAATCTTATGACAGCTTGTGCTTAACT 621
 Db 61 SerAlaGlyGluLeuThrLeuIleValLeuSerMetIleLeuAlaLeuProLeuSer 80
 Qy 622 TGTGTTGGGAATTAATGAGGAGTCTTACTGTTTATAGGAGGAGCACCTGTTGACTTTC 681
 Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 Qy 682 GAGACATACGACTTCTCAAAATGGGGAGCTCTTAAGTAATAGCGTCTGCTGATGACTG 741
 Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
 Qy 742 TTTACTATTGAGGTTTAAAGAAATTTATCTTTTCCAAATTTGCAATTTCAATCTTCCGTA 801
 Db 121 PheThrIleGluGlyPheIleValLeuSerPheSerAsnGlyAsnSerLeuLeuAlaVal 140
 Qy 802 CTGCTGCTGCAACGACTAATAGGGTAGCCAGACTCCGACGACCAACATCTACCGCTCT 861
 Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160

Qy 862 AATGGTACTATTTTCTAAAAACAGATCTTTTGTGTACTCAATATATGAGAGTCTCTCATTC 921
 Db 161 AsnGlyThrIleTyrSerIleThrAspLeuLeuLeuAsnAsnGluIlePheSerPhe 180
 Qy 922 TATAGTAAATTTAGTCTCTGGAGATGGGGAGCTATAGATCTAGAGCTTAACCGTTCAA 981
 Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaIleSerLeuThrValGln 200
 Qy 982 GGAATTAGCAAGCTTTGTCTTCCAAAGAAAATCTGCTCAAGCTGATGGGGAGCTTGT 1041
 Db 201 GlyIleSerIleValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
 Qy 1042 CAAGTACTCAAGCTTTCTGCTATGCTTAACAGAGCTCTCTATGCTTTGTAGCGAAT 1101
 Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
 Qy 1102 GTTCAGAGTAAAGGGGGAGGATGCTGCTGCTTCCAGATGGGAGCGAGGAGTGTCA 1161
 Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
 Qy 1162 TCATCTACTTCAACACAGAGATCCAGTAGTAGTATTTTCCAGAAAATCTGCGGTAGAGTTT 1221
 Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
 Qy 1222 GATGGAACTGAGCCCGAGTAGGAGGAGGATTTACTCTACGGGAACTGCTTCTCTG 1281
 Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
 Qy 1282 AATAATGGAAAACCTTGTCTCAACATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1341
 Db 301 AsnAsnGlyIleThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu 320
 Qy 1342 CAACCAACAGTGCAGAGCTTCTTAATACGAGTAGTAATTAATACGAGATGGAGGAGCTATC 1401
 Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
 Qy 1402 TTCTGTAAGAAATGGTGGCCAA---GCAGGATCCAAATACTCTGATCAGTTCTCTTGTAT 1458
 Db 341 PheCysIleAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 Qy 1459 GGAGAGGAGTAGTTTCTTTAGTAGCAATAGTCTGCTGGGAAAGGGGAGCTATTTAT 1518
 Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIleGlyAlaIleTyr 380
 Qy 1519 GCCAAAAGCTCTCGTTGCTAACTGTGCGCTGTACAAATTTTAAAGAAATATCTCTAAT 1578
 Db 381 AlaIleIleLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 Qy 1579 GATGTCGAGCGAGTTTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTTATGA 1638
 Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
 Qy 1639 GATATTTATTTGATGGGAATCTTAAAGAACAGCCAAAGAAATGCTCCGATGTTAAT 1698
 Db 421 AspIleIlePheAspGlyAsnLeuIleArgThrAlaIleGlyAsnAlaAlaAspValAsn 440
 Qy 1699 GCGCTAATCTGTCTCTCAAGCCATTTGATGGATCGGAGGAGGAAATAACGACATTA 1758
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460
 Qy 1759 AGAGCTAAGCAGGCGCATCAGATTTCTTTTAAATGATCCATCGAGATGGCAACGAAAT 1818
 Db 461 ArgAlaIleAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
 Qy 1819 AACCCAGCCAGCTCTTCCAAACTTCTAAATAATTAAGATGGTGAAGGATACACAGGG 1878
 Db 481 AsnGlnProAlaGlnSerSerGluProLeuIleAsnAspGlyGluGlyTyrThrGly 500
 Qy 1879 GATATTTGTTTGTCTTAATGGAGCAGTACTTTGTACCAAAATTTTACGATAGGCAAGA 1938
 Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
 Qy 1939 AGGATTGTTCTTCTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGAGGTGGG 1998

Db 521 ArgileValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGAGTACATGGATTTTGTAACTCCACACACACACACAG 2058
Db 541 SerLeuThrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTCCGCTAATCAGTTGATCACGCTTTTCCAACTGCTATGCTCTCTCTCTTTG 2118
Db 561 ProProAlaAlaGlnLeuLeuThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAAAATGAGTACGATTCGATCCCTACCAATCTCCAGCGGAGATTCATCCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCAATGCTAGCAACACTGCTGTTCTGTTTAACTAGTGGCCCTATCTTTTGG 2238
Db 601 AlaIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTTCTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTCTGAAATPACAGTTAGGAGCTAAGCCCGACCTAATGCCCCATCAGATTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGAAATCAGATGCTTAACTAGGTATCAAGAAAGCTGGAAGCTTGGTGGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
QY 2419 ACAGCAAAATATGCTCTTATCTACTGAAAGCTACATGGACTAAAGCTGGGTATAATCCT 2478
Db 681 ThrAlaAsnAsnGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnPro 700
QY 2479 GGGCTGAGCAGTAGCTCTTTGGTCCAAATAGTTATGGGATCCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
QY 2539 CGATCTGGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATGTCGAGGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740
QY 2599 TGGGTCTCTGAGTTTCGAAATTTCTTATCATGACCCGATGCTTTAGTTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPhePheTyHisAspArgAspAlaLeuGlyGlnGlyTy 760
QY 2659 CGGTATATTAGTGGGGTTATCTCTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAsnSerTyPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCATTTACGAGTATTTGGTAGATCTAAAGATTATAGTGTGCTGTTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAsn 800
QY 2779 CATCATGCTGCATAGGATCGGTTTATCTATCTACCCCAAGCTTTTATGTTGGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyIleSerThrLysGlnAlaLeuCysGlySerTy 820
QY 2839 TTGTTCCGAGATGCGTTTATCCGTCGTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTGCAGAGAGCGATGTTGTTGGGATATTAATGCTGCTGGCTGGAG 2958
Db 841 SerTyThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGAGCGGATTTACCGATTGTGATTACTTCCATCTAAGCTCTATTTGAAATGAGTTGGT 3018
Db 861 IleGlyValGlyLeuProIleThrProSerLysLeuTyThrLeuAsnGluLeuArg 880
QY 3019 CCTTTCGGCAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGAT 3078

Db 881 ProPheValGlnAlaGluPheSerTyAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGGCATTCAAGAGCGGACATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATATAGCTTTATGGCGGCTTATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTySerPheMetGlyAlaTyIleCys 940
QY 3199 GATGCTTATCGACCATCTCTGTACTGACACAACGCTCTATCCCATCAAGAGACATGG 3258
Db 941 AspAlaTyArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACACAGATCGCTTTCATTTAGCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyAla 980
QY 3319 TCTCTAAAGATATATAGAAATATATGGCCATCGGAAGATATGAGTATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyArgGlyHisGlyArgTyGluTyArgAspThrSer 1000
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 9
US-10-498-327-13
; Sequence 13, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-13

Alignment Scores:
Pred. NO.: 0 Length: 1013
Score: 5123.50 Matches: 979
Percent Similarity: 98.8% Conservative: 22
Best Local Similarity: 96.6% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-13 (1-1013)

QY 382 ATGCAAAAGCTTTTCCATAAGTCTCTTCTTCAATGATTCTAGCTTATTTCTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTySerCysCysSer 20
QY 442 TTAATATGGGGGGATATGCAGCAAAATCATGTTCTCTCAAGAAATTTACGATGGGAG 501
Db 21 LeuSerGlyGlyTyAlaAlaGluIleMetIleProGlnGlyIleTyAspGlyGlu 40
QY 502 ACGTTAATCTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTCAGGAGAGTTAAACATTAATAAATCTTCACAAATCTTATGAGCTTTTGCCTTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80

QY	622	TGTTTTGGGAACCTTATTAGGGAGTTTACTGTTTATAGGAGAGGACACTCGTTGACTTTC	681	Db	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyValIleThrThrLeu	460
Db	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100	QY	1759	AGAGCTAAAGCAGGGCATCAGATTCTCTTTAATGATCCATCGAGATGCAACAGCAAT	1818
QY	682	GAGAACATACGACTTCTACAAATGGGGCAGCTCTAAAGTAATAGCGCTCTGATGACTG	741	Db	461	ArgAlaIysAlaGlyHisGlnIleLeuPheAsnAspPheIleGluMetAlaAsnGlyAsn	480
Db	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120	QY	1819	AACCAAGCAGCGCAGCTCTTCCAAACTTCTAAAAATTAACGATGGTGAAGGATACACAGG	1878
QY	742	TTTATCTATTGAGGGTTTTAAGAATTATCTCTTTTCCAATGTCATTTCACTTCCGCTA	801	Db	481	AsnGlnProAlaGlnSerSerGluProLeuIleAsnAspGlyGluGlyThrGly	500
Db	121	PheThrIleGluGlyPheIysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140	QY	1879	GATATTGTTTGTCTAATGGAAGCAGTACTTTGTACCAAAAATGTTACGATAGAGCAAGCA	1938
QY	802	CTGCGCTGCGCAACGACTTAATAGGCTAGCCAGACTCCGACGACACATCTACACCGTCT	861	Db	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
Db	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160	QY	1939	AGGATTGTCTCTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG	1998
QY	862	AATGGTACTATTATTCTAAAACAGATCTTTTGTACTCAATAATCAGAAGTTCTCATTC	921	Db	521	ArgIleValLeuArgGlnIlyAlaIysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
Db	161	AsnGlyThrIleTyrSerIysThrAspLeuLeuLeuLeuAsnGluIysPheSerPhe	180	QY	1999	AGTCTGTATATCGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACCAACACAG	2058
QY	922	TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGACTTTAAGCTTCAA	981	Db	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
Db	181	TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaIysSerLeuThrValGln	200	QY	2059	CTCTCCCGCTAATCAGTTGATCAGCTTTCGAATCTGCATTTGCTCTTTCTTTCTTTG	2118
QY	982	GGNATTAGCAAGCTTTGTCTCTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT	1041	Db	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu	580
Db	201	GlyIleSerIysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyIlyAlaCys	220	QY	2119	TTAGCAAAACAATGCAGTTACGAATCTCTTACCCTCCAGCGCAAGATTTCTCATCT	2178
QY	1042	CAAGTAGTCACCAAGTTTCTCTGCTATGGCTAAACGAGGCTCTTATTCGCTTTGAGCGAAT	1101	Db	581	LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro	600
Db	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240	QY	2179	GCAGTCATTGGTAGCAACAACCTCTGCTTCTGTACAAATTAGTGGGCTATCTTTTGTAG	2238
QY	1102	GTTTCAGGAGTAAGAGGGGAGGATGCTGCTGTTCCAGATGGCGAGGAGGTGCA	1161	Db	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
Db	241	ValAlaGlyValArgGlyGlyLeuAlaAlaValGlnAspGlyGlnGlnGlyValSer	260	QY	2239	GATTTCCGATGATACAGCTTATGATAGTATGATGGCTAGTGTCTTAATCAAAAATCAAT	2298
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACTCGCGGTAGAGTTT	1221	Db	621	AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnIlyAsp	640
Db	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280	QY	2299	GTCTCGAAATACAGTTAGGACTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA	2358
QY	1222	GATGGGAACCTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCTTCTCCTG	1281	Db	641	ValLeuIysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
Db	281	AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300	QY	2359	GGGAATGAGATGCTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGGTGGGATCTTAAT	2418
QY	1282	AATAATGGAAAAACCTTTGTTCTCAACAATGTTGCTTCTCTGTTTACATGCTGCTAAG	1341	Db	661	GlyAsnGluMetProIysTyrGlyTyrGlnGlySerTyrIysLeuAlaIleTyrAspProAsn	680
Db	301	AsnAsnGlyIysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu	320	QY	2419	ACAGCAAAATAATGGTCTCTTATCTCTGAAAGCTACATGGAATAAACTGGGTATATCT	2478
QY	1342	CAACCAACAAGTGGACAGGCTTCTAATACGAGTAATAATACGAGATGAGGAGCTATC	1401	Db	681	ThrAlaAsnAsnGlyProTyrThrLeuIysAlaThrTriphThrIysThrGlyTyrAsnPro	700
Db	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyIlyAlaIle	340	QY	2479	GGGCTGAGCGAGTAGTCTTTGGTTCGAAATAGTTATGGGATCCATTTTAGATATA	2538
QY	1402	TTCTGTAAAGATGGTGGCAAA---GCAGGATCCAATAACTCTGGATCAGTTTCTCTTGAT	1458	Db	701	GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle	720
Db	341	PheCysIysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360	QY	2539	CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGGCGCTCTTATTGTCAGGAGTAA	2598
QY	1459	GGAGAGGAGTAGTATTTCTTTAGTAGCAATGTAGCTCTGGGAAGGGGAGCTATTAT	1518	Db	721	ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu	740
Db	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIysGlyIlyAlaIleTyr	380	QY	2599	TGGGTTTCTGGAGTTTCGAAATTTCTTCTATCATGACCGCATGCTTTAGTCTCAGGAGTAT	2658
QY	1519	GCCAAAAGCTCTCGGTGCTAACTGCGCCCTGTACAAATTTTAAAGGAATATCGCTAAT	1578	Db	741	TrpValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr	760
Db	381	AlaIysIysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400	QY	2659	CGGTATATTAGTGGGGTTATCTCTTAGGAGCAACTCTACTTTGGATCATCTCATCTTT	2718
QY	1579	GATCGTCGAGGATTTTATAGGAGAATCTGAGAGCTCAGTTTATCTGCTGATTATGGA	1638	Db	761	ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe	780
Db	401	AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly	420	QY	2719	GGTCTAGCAATTTACCGAAGTATTGGTAGATCTAAAGATTATGATGTGTGCTGTCCAAT	2778
QY	1639	GATATTATTTTCGATGGGAATCTTTAAAAAGAACAGCCAAAGAGAAATGCTCGCGATTTAAT	1698	Db	781	GlyLeuAlaPheThrGluValPheGlyArgSerIysAspTyrValValCysArgSerAsn	800
Db	421	AspIleIlePheAspGlyAsnLeuIysArgThrAlaIysGluAsnAlaAlaAspValAsn	440	QY	2779	CATCATCTGTGCATAGGATCCGTTTATCTATCTACCCAAACAGCTTTATGTGGATCTTAT	2838
QY	1699	GGCGTAATGTGTCTCTCAAGCCATTTTCGATGGGATCGGGAGGAAAAATACACATTA	1758				

Db 801 HisHlaCyAlleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTCCGAGATGCGTTTATCCGTCTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2898 TCATATACATTCGACAGAGAGCGATGTCGTTGGGATAATACTGCTCGCTGGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGAGCGGATACCGATTGTGATTACTTCCATCTAAGCTCTATTGAAATGAGTTGGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy 3019 CTTTTCGTCGAGCTGAGTTTCTTATCCCGATCATGAATCTTTTACAGAGAAAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGCATTCACAGCGGACATCTCCTAAATCTATCATGTTCTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTCATCATGTTCTAGTACACATCTTAATAATATAGCTTTATGCGCGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCGCACCATCTCTGGTACTGAGACACCGCTCCTATCCCATCAAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATGCTTCATTTAGCAGACATGGAGTGTGGTTAGAGGATCATGTATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy 3319 TCTTAAACAAGTAAATAGAGTATATGCCATGGAAGATAGATATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGCTATGTTGAGTCGAGGAAGTAGAGTCCGTTCT 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 10

US-10-498-327-93
; Sequence 93, Application US/10498327
; Publication NO. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 93
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-93

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5123.50 Matches: 979
Percent Similarity: 98.8% Conservative: 22
Best Local Similarity: 96.6% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-93 (1-1013)

Qy 382 ATGCAACGCTTTCATAGTTCTTCTTCAATGATCTTAGCTTATTTCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
Qy 442 TTAATGGGGGGGATATGACAGAAATCATGTTCTCAAGGAATTTACGATGGGGAG 501
Db 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
Qy 502 ACGTTAACTGATCATTTCCCTATATCTGTTATAGAGATCCGAGTGGAGCTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTCAGGAGAGTTAAACATTAATAAATCTTGACAAATCTATTCGAGCTTCCCTTAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer 80
Qy 622 TGTTTGGGAACCTTATTAGGAGTTTACTGTTTGGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTG 741
Db 101 GluAsnIleAArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTACTATTGAGGCTTTTAAAGAATTATCCTTTTCCAAATTCGAATTCATTACTTCCCTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTGCTGCTGCAACGAGCTAATAAGGTAGCGAGCTATAGATGCTTAAGCTTAACGGTTCAA 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer 160
Qy 862 AATGCTACTATTATTCTTAAACAGACTTTTGTGTACTTGTACTCAATATAGAGAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
Qy 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
Qy 982 GGAATTAGCAAGCTTTGTGCTTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTCACGAGTTTCTCTGATATGCTTAACGAGGCTCTTATTTGCTTTGACGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy 1102 GTTCGAGGATAGAGGGGGGAGTTGCTGCTGTTTCCAGGATGGCGACGGAGGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlyValSer 260
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTCGCGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Qy 1222 GATGGAAACGTACCCGAGTAGGAGGAGGATTTACTCTACGGGAAACGTTGCTTTCCGTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Qy 1282 AATAATGGAAAAACCTTGTCTCAACAATGTTGCTTCTCTGTTTACATTCGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrLeuAlaAlaGlu 320
Qy 1342 CAACCAACAAGTGGACAGCTTCTAATACAGTAATAATAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyAlaIle 340
Qy 1402 TTCTGTAAAGATGGTGGCGAA---GCAGATCCAATAACTCTGATACGATTTCTTTGAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
Qy 1459 GGAGAGGAGTAGTTTCTTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518

Db 361 GlyGluGlyValValPheSerSerSerValAlaAlaGlyGlyGlyAlaLeuTyr 380
QY 1519 GCCAAAGCTCTCGGTCTCTAACTGTGCGCCCTGTACAAATTTTAAAGGAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAlaCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
QY 1579 GATGGTGGAGCAATTTATTTAGGAGAACTCTGGAGAGCTCAGTTTATCTCTGCTGATTATGA 1638
Db 401 AspGlyGlyAlaLeuTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTATTTCGATCGGGAATCTTAAAGAAACAGCCAAAGAGAGAAATCTCCGATGTTAAT 1698
Db 421 AspileilePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGTAACCTGTGCTCAACAGCCATTTCGATGGGATCGGGAGGGAAATAACAGCAATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGCCAAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACCCAGCAGCGAGCTTCCAAACTTCTAAATAATTAAACGATGGTGAAGGATACACAGGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyGlyTyrThrGly 500
QY 1879 GATATTGTTTTCTTAATGGAAGCAGTACTTTCTACCAAAATTTTACGATPAGAGCAAGGA 1938
Db 501 AspileValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGAATTGTTCTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGCTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGAAGCTGGAGTACATGGGATTTGTAACTCCAAACACACACACACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTCGCCCTAATCAGTTGATCAGCTTTCCAACTTCCAACTTGTCTCTTTCTTTG 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAAAATCAGAGTTAGCAATCTCTCCACCAATCTCCAGGCGAAGATTTCTCATCCT 2178
Db 581 LeuAlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCAATTCGTAGCACAACTGCTGTTCTGTACAAATTAGTGGCCCTATCTTTTGTAG 2238
Db 601 AlaIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTGGATGATACAGCTTATGATPAGGTATGATTGGCTAGGTCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAspThrAlaLysAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTGAAATTAAGTTAGGACCTAAGCCCGCCAGCTAATGCCCATCAGATTGCACTTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGATCAGATGCTTAAGTATGGCTATCAAGGAGCTGGAAGCTTGGTGGATCCTTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaIlePheProAsn 680
QY 2419 ACAGCAAAATATGCTCTTACTCTGAAAGCTACATGGACTAAAAGCTGGGTATAATCCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrIlePheThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCTCAGCAGTAGCTTCTTTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspile 720
QY 2539 CGATCTCGCAATTCAGCAATTCAGAGAGTGTGGATGGCGCTCTTATTGTGAGGATTA 2598

Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGAGATTTTCGAATTTCTCATATGACCGCGATGCTTTTAGTTCAGGATAT 2658
Db 741 trpValSerGlyValSerAsnPheThrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCTTTAGGAGCAAACTCTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
QY 2719 GGTCTAGCAATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
QY 2779 CATCATGCTTGCATAGGATCCGTTTATCTATCTACCCAAAGCTTTTATGTGCATCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY 2839 TTGTTTCGGAGATCGGTTTATCCGCTAGCTACGCTTGGGAATCAGCATATGAAAAAC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
QY 2899 TCATATACATTTTCGAGAGAGAGCGATGTTCTGTTGGGATAATACTGTCTGGCTCGAGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTyrAspAsnAsnCysLeuValGlyGlu 860
QY 2959 ATTGGAGCGGATTAACGATGTTACTTCCATCTAAGCTCTTATTTGAATGATGTCGT 3018
Db 861 IleGlyValGlyLeuProIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
QY 3019 CTTTTCGTCACCAAGCTGAGTTTCTTATGCGCATCATGAACTTTTACAGAGGAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGlyAsp 900
QY 3079 CAAGCTCGGCGCTTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCTCTAAATATAGCTTTTATGCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCGCACCACTCTCTGCTAGACAGCAACGCTCTCTATCCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGCAACACATGGAGTTGTGGTTAGAGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAATATAGAGTATATGGCATGGAAGATATGATGATATCGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGGCTATGTTGAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 11

US-10-498-327-1
; Sequence 1, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498.327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2


```
Db      641  ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660
QY      2359  GGAATGAGATCCCTAAGTATGGCTATACAGAGCTGAAGCTTGGCTGGATCCTAAT 2418
Db      661  GlyAenGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAen 680
QY      2419  ACAGCAAAATAAGTCCCTATATCTCTGAAGCTACATGAGCTAAAGCTGGTATATCTCT 2478
Db      681  ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro 700
QY      2479  CGGCTCAGCAGTATCTCTTTGGTTCCAAATAGTTATGGGATCCATTTTATAGATATA 2538
Db      701  GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
QY      2539  CGATCTGCCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTTCTCGAGATTA 2598
Db      721  ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY      2599  TGGGTTCTGGAGTTTCGAATTTCTCTATCATGACCGCATGCTTTAGTTCAGGATAT 2658
Db      741  TrpValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY      2659  CGGTATATTAGTGGGGTTTATCTCTTAGGAGCAAACTCTCTACTTTGGATCATCGATGTT 2718
Db      761  ArgTyrIleSerGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
QY      2719  GGTCTAGCATTTACCAAGTATTTGGTAGATCTAAAGATTTAGTAGTGTCTTCCAT 2778
Db      781  GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
QY      2779  CATCATGCTGCATAGATCCGTTTATCTCTATCTACCCCAACAGCTTTATGTGATCCTAT 2838
Db      801  HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
QY      2839  TTGTTCCGAGATCGGTTTATCCGCTAGCTACGGGTTTGGGAATCAGCATATGAAAAAC 2898
Db      821  LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
QY      2899  TCATATACATTTGCAGAGAGCGATGTTCTGTTGGGATATTAATCTGTCTGGCTGGAGAG 2958
Db      841  SerTyrThrPheAlaGluGluSerAspValArgTrpAspAenAenCysLeuValGlyGlu 860
QY      2959  ATTGGAGCGGATACCGATTGTAATCTCTCAATCTTAAGCTCTATTGATGATGTCGT 3018
Db      861  IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
QY      3019  CTTTCTGTCAGCTCAGATTTCTTATGCGCATCATGAATCTTTTACAGAGGAAGCGAT 3078
Db      881  ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluAsp 900
QY      3079  CAAGCTCGGCGATTCAGAGCGGACATCTCTCTAAATATATCAGTTCTCTGTTGGAGTGAAG 3138
Db      901  GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
QY      3139  TTTGATCGATGTTCTAGTACACATCTCTAAATATAGCTTTATGGCGCTTATATCTGT 3198
Db      921  PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940
QY      3199  GATGCTTATCGCACCATCTCTGTTACTGAGACACAGCTCTCTATCCCATCAAGAGACATGG 3258
Db      941  AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
QY      3259  ACAACAGATCCCTTTTATAGCAAGACATGGAGTGTGTTAGAGATCTATGTATGCT 3318
Db      961  ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY      3319  TCTTAACAAGTATATAGAATATATGGCCATGGAAGATATGATATCGAGTCTCTCT 3378
Db      981  SerLeuThrSerAenIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY      3379  CGAGGCTATGTTGAGTGCAGGAGTAGAGTCCGGTTC 3417
Db      1001  ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
```

```
RESULT 12
US-10-498-327-3
; Sequence 3, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Ratti, Guido
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 02441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-3
```

```
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5121.50 Matches: 978
Percent Similarity: 98.8% Conservative: 23
Best Local Similarity: 96.5% Mismatches: 11
Query Match: 65.0% Indels: 1
DB: Gaps: 1
```

```
US-10-701-844-1 (1-4435) x US-10-498-327-3 (1-1013)
QY      382  ATGCAAACTCTTCCATAAGTTCTTCTTCAATGATTCTAGCTTATTCTTCTGCTCT 441
Db      1  MetGlnThrSerPheHisLysPhePheLeuSerMetLeuLeuAlaTyrSerCysSer 20
QY      442  TTAATATGGGGGGATATGCAGCAGAAATCATGGTTCTCAAGGAATTTACGATGGGAG 501
Db      21  LeuSerGlyGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
QY      502  AGTTAACTGTATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTTTT 561
Db      41  ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY      562  TCTGCAGGAGACTTAACATTAATAAATCTTGACAAATCTTGCAGCTTTGCTTTAAGT 621
Db      61  SerAlaGlyGluLeuThrLeuLysAenLeuAspAenSerLeuAlaLeuProLeuSer 80
QY      622  TGTTTTGGGAACCTTATTAGGGAGTTTACTGTTTTATAGGAGAGACACACTCGTTGACTTC 681
Db      81  CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY      682  GAGAAATACGACTTCTCAAAATGGGCGAGCTCTAAGTAATAGCCGCTGCTGATGACTG 741
Db      101  GluAenIleArgThrSerThrAenGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
QY      742  TTTTACTATTGAGGGTTTAAAGAAATATCTCTTTTCAATTTGCAATTCATTACTTCCGCTA 801
Db      121  PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
QY      802  CTCCTCTGTCACAGACTTAATAAGGTTAGCCAGACTCCGACGACCAACATCTACACCGTCT 861
Db      141  LeuProAlaAlaThrThrAenAenGlySerGlnThrProThrThrThrSerThrProSer 160
QY      862  AATGGTACTATTATTCTTAAACAGATCTTTTGTGTTACTCAATTAATGAGAAGTTCTCATTC 921
Db      161  AenGlyThrIleTyrSerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180
QY      922  TATAGTAATTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAA 981
Db      181  TyrSerAenLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY      982  GGAATTAGCAAGCTTTTGTGTTCTTCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
```

Db 201 GlylleSerLysLeuCysValPheGlnGluSerThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTCACAGTTCTCTGCTATAGCTAACGAGGCTCTTATGCTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240
Qy 1102 GTTCAGAGTAAGAGGGGAGGATTGCTGCTTCAAGTGGCAGCAGGAGTCACA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyLeuAlaAlaValGlnAspGlyGlnGlyValSer 260
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTCAGTTCCTCAGAAATACCTGCGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
Qy 1222 GATCGGAACGTAGCCCGAGTAGGAGGAGATTACTCTCCAGGAACGTTGCTTCTCTG 1281
Db 281 AspGlyAenValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAenValAlaPheLeu 300
Qy 1282 AATAATGGAACCTTGTCTCTCAACATGTTGCTTCTCTGTTTACATGCTGCTAAG 1341
Db 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrLeuAlaAlaGlu 320
Qy 1342 CAACCAACAAGTGGACAGGCTTCTAATACGAGTAATAATTACGAGAGATGGAGGACTATC 1401
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAenThrSerAenThrGlyAenValAlaIle 340
Qy 1402 TTCTCTAAGAAATGGTGGCAA---GCAGATCCAAATACTCGGATCAGTTTCCTTGAAT 1458
Db 341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAen 360
Qy 1458 CGAGAGGAGTAGTTTCTTTAGTAGCAATGCTGCTGGGAAAGGGGAGCTATTAT 1518
Db 361 GlyGluGlyValAlaPhePheSerSerAenValAlaAlaGlyLysGlyAlaIleTyr 380
Qy 1519 GCCAAAAGCTCTCGGTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATACGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
Qy 1579 GATGTGGAGCGAATTTATTAGAGAACTCGAGAGCTCAGTTTATCTGCTGATATGGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
Qy 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAACTGCTCCGATTTAAT 1698
Db 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440
Qy 1699 GCGCTAACTGCTCTCAAGCCATTTCGATGGGATCGGAGGAGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
Qy 1759 AGAGCTAAAGCGGATCAGATTCTCTTTAATGATCCCATCGAGATGCGAAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAspProIleGluMetAlaAenGlyAen 480
Qy 1819 AACCGCCAGCGAGCTTCTCCAAACTTCTAAAAATTAACGATGGTGAAGATACACAGG 1878
Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAspGlyGlyIleThrGly 500
Qy 1879 GATATTGTTTTGCTTAATGGAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGA 1938
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyrGlnAenValThrIleGluGlnGly 520
Qy 1939 AGGATTGTTCTTCGTGAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540
Qy 1999 ACTCTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCCAACCAACCAACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAenPheValThrProGlnProGlnGln 560
Qy 2059 CCTCTGCGCGCTAAATCAGTTGATCAGCTTTTCCAATCTGCATTTGTCTCTTCTTTG 2118

Db 561 ProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeu 580
Qy 2119 TTACAAACATGCAAGTTACGAATCCTCTACCAATCCTCCAGCGCAGATTCATCATCT 2178
Db 581 LeuAlaAenAenAlaValThrAenProProThrAenProProAlaGlnAspSerHisPro 600
Qy 2179 GCAGTCATTGTTAGCACAACTGCTGTTCTGTTACAAATAGTGGGCCCTATCTTTTTCAG 2238
Db 601 AlalleIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATTGGCTAGTCTTAATCAAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAenGlnLysIleAen 640
Qy 2299 GTCTGAAATTTACAGTTAGGCACTAAGCCCCACCTAATGCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAenAlaProSerAspLeuThrLeu 660
Qy 2359 GCGAATGAGATGCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTTGCCTGGATCCTAAT 2418
Db 661 GlyAenGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaIleTrpAspProAen 680
Qy 2419 ACAGCAATAATGCTCTTATCTACTCTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT 2478
Db 681 ThrAlaAenAenGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAenPro 700
Qy 2479 GGGCTGAGCAGTAGCTTCTTGGTTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIle 720
Qy 2539 CGATCTGGCATTCAGCAATTCAGCAAGTGTGAGTGGCGCTCTTATTTGTCAGAGATA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyLysSerTyrCysArgGlyLeu 740
Qy 2599 TGGGTTTCTGAGGATTCGAATTTCTCTATCATGACCCGATGCTTTAGTTCAGGAGAT 2658
Db 741 TrpValSerGlyValSerAenPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGGACAAACTCCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAenSerTyrPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGATGTGCTGTTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAen 800
Qy 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTATCAACCAACAGCTTTATGTCATCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTCCGAGATGCGTTTATCCGTCTAGCTACCGGTTTGGGAATCAGCATATGAAACC 2898
Db 821 LeuPheGlyAspAlaPheLeuArgAlaSerTyrGlyPheGlyAenGlnHisMetLysThr 840
Qy 2899 TCATATACATTTGCAGAGGAGCGATGCTCGTTGGGATAATACTGCTGCTGCTGAGAG 2958
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAenAenCysLeuValGlyGlu 860
Qy 2959 ATTGAGCGGAGATTACCGATTGTGATTACTCCATCTAAGCTCTATTTGAATGAGTTGCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAenGluLeuArg 880
Qy 3019 CTTTCTGTCGAAAGCTGAGTTTCTTATCCGATCATGATCTTTTACAGAGAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGGCAATTCAGAGGAGCATCTCTCTAAATCTATCAGTTCTCTGTTGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCGATGTTCTAGTACATCTCTAATAATATAGCTTTATGGCGGCTTATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAenLysTyrSerPheMetGlyAlaTyrIleCys 940

QY 1906 ACTTTGTACAAAATGTTACATAGACGACGAGGATTGTTCTTCTGTGAAAGCGCAAAA 1965
DB ThrLeuTyrGlnAenValThrIleGluGlnGlyArgIleValLeuArgGluLeuAlaLys 500
QY 1966 TTATCAGTGAATCTCTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGCAGTACA 2025
DB LeuSerValAenSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThr 520
QY 2026 TGGGATTTGTAACTCCACAAACCCACACAGAGCTCTCGCGCTTAATCAGTTCATCAGC 2085
DB TrpAspPheValThrProGlnProGlnGlnProProAlaAlaAenGlnLeuIleThr 540
QY 2086 CTTTCCAAATCTGCATTTGCTCTTTCTTTCTTTGTTAGCAAAATGCGAGTTACGATCCT 2145
DB LeuSerAenLeuHisLeuSerLeuSerSerLeuLeuAlaAenAenAlaValThrAenPro 560
QY 2146 CCTACCAATCTCCAGCGCAAGATTCTCATCTCGCAGTCATTGGTAGCACAACTGCTGGT 2205
DB ProThrAenProProAlaGlnAaspSerHisProAlaValIleGlySerThrAlaGly 580
QY 2206 TCTGTTACAAATTAGTGGCCTATCTTTTTCAGGATTTGGATGATACAGCTTATCATAGG 2265
DB SerValThrIleSerGlyProIlePhePheGluAepLeuAepThrAlaTyrAaspArg 600
QY 2266 TATGATTCGCTAGGTTCTTAATCAAAAATCAATGTCCTGAAATTTACAGTTAGGAGCTAAG 2325
DB TyrAaspTrpLeuGlySerAenGlnIleAenValLeuLysLeuGlnLeuGlyThrLys 620
QY 2326 CCCCAGCTAATGCCCATCAGATTGACTCTAGGGAATGAGTCCCTAAGTAGGCTAT 2385
DB ProProAlaAenAlaProSerAaspLeuThrLeuGlyAenGluMetProLysTyrGlyTyr 640
QY 2386 CAAGAACTGGAAGCTTCGCTGGGATCTTAATACAGCAAAATAATGGTCCTTATCTCTG 2445
DB GlnGlySerTrpLysLeuAlaTrpAspProAenThrAlaAenAenGlyProTyrThrLeu 660
QY 2446 AAAGCTACATGACTAAACTGGGTATTAATCTCGGCTCGAGCGAGTAGCTCTTTGGTT 2505
DB LysAlaThrTrpThrLysThrGlyTyrAenProGlyProGluAtrValAlaSerLeuVal 680
QY 2506 CCNAATAGTTATGGGGATCCATTTTATAGATATACGATCTGGCATTCAGCAATTCAGCA 2565
DB ProAenSerLeuTrpGlySerIleLeuAaspIleArgSerAlaHisSerAlaIleGlnAla 700
QY 2566 AGTGTGATGGCGCTCTTATGTTCGAGATTTATGGTTCTTCGAGTTTCGAAATTTCTTC 2625
DB SerValAaspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAenPhePhe 720
QY 2626 TATCATGACCGGATGCTTTAGGTACGGATATCGGTATATAGTGGGGTTATTCCTTA 2685
DB TyrHisAaspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeu 740
QY 2686 GGAGCAAACTCTACTTTGGATCATCGATGTTTGTCTAGCATTTACCGAGTATTTGCT 2745
DB GlyAlaAenSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760
QY 2746 AGATCTAAAGATTATGTAGTGTGCTTCCAAATCATCATGCTTGATAGGATCCGTTTAT 2805
DB ArgSerLysAaspTyrValValCysArgSerAenHisHisAlaCysIleGlySerValTyr 780
QY 2806 CTATCTACCCAAACAGCTTTATGTGATCTCTATTTGTTTCGGAGATGCGTTTATCCGTCT 2865
DB LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAaspAlaPheIleArgAla 800
QY 2866 AGCTACGGTTTGGGATCAGCATATGAAACCTCATATACATTTTGCAGAGGAGCGCAT 2925
DB SerTyrGlyPheGlyAenGlnHisMetLysThrSerTyrThrPheAlaGlnGluSerAasp 820
QY 2926 GTTCGTGGGATAAATAACTGCTCGCTGGAGAGATTGGAGCGGGATTACCGATTGTGATT 2985
DB ValAsgTrpAaspAenCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
QY 2986 ACTCCATCTAAGCTCTATTGTAATGAGTTGCTCTTCCTCGTCCAAAGCTGAGTTCTTAT 3045

DB 841 ThrProSerLysLeuTyrLeuAenGluLeuArgProPheValGlnAlaGluPheSerTyr 860
QY 3046 GCCGATCATGAATCTTTTACAGAGGAAAGCGCATCAAGCTCGGCGATTCAAGAGCGGACAT 3105
DB AlaAaspHisGluSerPheThrGluGluGlyAaspGlnAlaArgAlaPheLysSerGlyHis 880
QY 3106 CTCCTAAATCTATCAGTTCCTGTTGGAGTGAGTTGATCGATGCTTCTAGTACACATCCT 3165
DB LeuLeuAenLeuSerValProValGlyValLysPheAaspArgCysSerSerThrHisPro 900
QY 3166 AATAAATATAGCTTTATGGCGCTTATATCTGTCATGCTTATCGCACCATCTCTGGTACT 3225
DB AenLysTyrSerPheMetAlaAlaTyrIleCysAaspAlaTyrArgThrIleSerGlyThr 920
QY 3226 GAGCAACGCTCTCTATCCCATCAAGACATGGAACAACAGATGCTTTCATTTTAGCAAGA 3285
DB GluThrThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAlaArg 940
QY 3286 CATCGAGTTGCTTAGAGATCTATGATGCTTCTCTAACAGTAATATAGAGTATAT 3345
DB HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAenIleGluValTyr 960
QY 3346 GGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTGAGTGCAGGAAGT 3405
DB GlyHisGlyArgTyrGluTyrArgAaspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
QY 3406 AGAGTCCCGGTTTC 3417
DB ArgValArgPhe 984
RESULT 14
US-10-766-711-43
; Sequence 43, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-43
Alignment Scores:
Pred. No.: 0 Length: 984
Score: 5116.00 Matches: 984
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.9% Indels: 0
DB: 4 Gaps: 0
US-10-701-844-1 (1-4435) x US-10-766-711-43 (1-984)
QY 466 GAAATCATGTTCTCTCAAGGAATTTACATGGGAGAGCTTAACGTGATCATTTCCCTAT 525
DB 1 GluIleMetValProGlnGlyTyrAspGlyGluThrLeuThrValSerPheProTyr 20
QY 526 ACTGTTATAGGAGATCCAGTGGGAGTACTGTTTTTTCTGCGAGGAGCTTAACATTAAA 585
DB 21 ThrValIleGlyAaspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAATCTTATTGACAGCTTGGCTTTTAAAGTTGTTTGGGAACTTATTAGGAGT 645
DB 41 AenLeuAaspAenSerIleAlaLeuProLeuSerCysPheGlyAenLeuLeuGlySer 60

646 TTTACTGTTTGGGAGGACACTCGTTGACTTTTCGAGACATACGGACTTCTACAAT 705
Db PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
706 GGGGAGCTCTAAGTAATAGCCCTGCTGATGACTGTTTACTATTATGAGGGTTTTAAAGAA 765
Db 81 GlyAlaIleLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheGlyGlu 100
766 TTTACTCTTTTCCAATTGCAATTCATTACTGTCGGTACTGCTGCTGCGCAACGACTTAATAG 825
Db 101 LeuSerPheSerAsnCyAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
826 GGTAGCCAGACTCCGACGACACACTACACCGTCTAATGGTACTATTATTCTAAACA 885
Db 121 GlySerGlnThrProThrThrThrThrProSerAsnGlyThrIleIleIleIleIleIle 140
886 GATCTTTTCTTACTCAATAATGAGAAGTTCTCATCTATAGTAATTTAGTCTCTGGAGAT 945
Db 141 AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTySerAsnLeuValSerGlyAsp 160
946 GGGGAGCTATAGACTTAAGACTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTC 1005
Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCyValPhe 180
1006 CAAGAAATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCAACAGTTTCTCTGCT 1065
Db 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCySerGlnValValThrSerPheSerAla 200
1066 ATGGCTAACGAGCTCTATTGCTTTGTAGCAATGTTGCAGAGTAAGAGGGGAGGG 1125
Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGly 220
1126 ATTGCTGCTTTCAGGATGGCAGCAGGGAGTGTCAATCTACTTCAACAGAGATCCA 1185
Db 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerSerThrSerThrGluAspPro 240
1186 GTAGTAAGTTTTCCGAAATACTCGGCTAGATTGATGGGACCTAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
1246 GAGGAGATTACTCTTACGGGACGTTGCTTTCTCGAATAATGGAACCTTTGTTTCTC 1305
Db 261 GlyGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 280
1306 AACAAATGTTGCTCTCTGTTTACATTGCTGTGCTTAAGCAACCAAGTGACAGGCTTCT 1365
Db 281 AsnAsnValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
1366 AATACGAGTAATTAATACGAGATGGAGGAGCTATCTTCTGTAAGATGGTGCGCAGCA 1425
Db 301 AsnThrSerAsnAsnTyGlyAspGlyGlyAlaIlePheCyLysAsnGlyAlaGlnAla 320
1426 GGATCCAATAACTCTGATCAGTTTCTTTGATGGAGAGGAGTAGTTTTCTTTAGTAGC 1485
Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSer 340
1486 AATGTAGCTCTCGGAAAGGGGAGCTATTATTATGCCAAAAGCTCTCGGTTGCTAACTGT 1545
Db 341 AsnValAlaAlaGlyLysGlyGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCys 360
1546 GGCCTGTACAATTTTAAGGAATATCGCTAATGATGGTGGAGGATTTATTATTAGGAA 1605
Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyLeuGlyGlu 380
1606 TCTGGAGACTCAGTTTATCTGCTGATTATGGAGATATTATTTCCGATGGATCTTAAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsnLeuLys 400
1666 AGAACACCCAAAGAGATGCTGCGATGTTAATGGCGTAACTGTGCTCTCACAGCCATT 1725
Db 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420

1726 TCGATGGATCGGAGGAGAAATAACGACATTAAAGACTAAAGCAGCGCATCAGATTCTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
1786 TTTAATGATCCATCGAGATGCGCAACCGAAATAACAGCCAGCGCAGCTCTCCAAACTT 1845
Db 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
1846 CTAAAAATTAACGATGGTGAAGGATACACAGGGGATATTGTTTTTGTCTAATCGAAGCAGT 1905
Db 461 LeuLysIleAsnAspGlyGlyGlyTyThrGlyAspIleValPheAlaAsnGlySerSer 480
1906 ACTTTGTACAAAATGTTACGATAGCAAGCAAGGATGTTCTTCGCGAAAAGGCAAAA 1965
Db 481 ThrLeuTyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
1966 TTATCAGTGAATTTCTTAAGTCAGACAGCTGGAGTCTGTATATGGAAGCTGGAGTACA 2025
Db 501 LeuSerValAsnSerLeuSerGlnThrGlySerLeuTyMetGluAlaGlySerThr 520
2026 TGGGATTTGTAACTCCACCAACCAACCAACAGCTCTCTGCGCTAATCAGTTGATCAGC 2085
Db 521 TrpAspPheValThrProGlnProGlnProAlaAlaAsnGlnLeuIleThr 540
2086 CTTTCCAATCTGCATTGTTCTCTTTCTTTGTAGCAAAACAATGCAAGTTGATCAGTCT 2145
Db 541 LeuSerAsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAlaValThrAsnPro 560
2146 CTACCAATCTCAGCGCAAGATTCTCATCTGCACTGTCATTTGTTAGTACCAACTGCTGCT 2205
Db 561 ProThrAsnProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGly 580
2206 TCTGTTACAAATTAGTGGGCTATCTTTTGTAGGATTTTCGATGATACAGCTTATGATAGG 2265
Db 581 SerValThrIleSerGlyProIlePhePheGluAspLysAspThrAlaTyAspArg 600
2266 TATGATGGCTAGGTTCTTAATCAAAAAATCAATGTCTCGAAATACAGTTAGGACTAAG 2325
Db 601 TyAspTrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlyThrLys 620
2326 CCCCAGCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCTAAGTATGCTAT 2385
Db 621 ProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyGlyTy 640
2386 CAAGGAGCTCGAAGCTTCGCTGGGATCTTAATACAGCAAAATAATGCTCTTATCTCTG 2445
Db 641 GlnGlySerTrpLysLeuAlaIlePheProAsnThrAlaAsnAsnGlyProTyThrLeu 660
2446 AAAGCTACATGGACTAAAACTGGGTATAATCTCGGCCCTGAGCGAGTAGCTCTTTGGTT 2505
Db 661 LysAlaThrTrpThrLysThrGlyTyAsnProGlyProGluArgValAlaSerLeuVal 680
2506 CCAATAGTTTATGGGATCCATTTTAGATATACGATCTCGGCATTCAGCAATTCAGCA 2565
Db 681 ProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700
2566 AGTGTGGATGGGCTCTTATTGTCGAGGATATGCTTCTGAGTTCGAAATTTCTTCT 2625
Db 701 SerValAspGlyArgSerTyCyAspGlyLeuTrpValSerGlyValSerAsnPhePhe 720
2626 TATCATGACCGCATCTTTTAGCTCAGGATATCGGTATATATTAGTGGGGTATTCTCTTA 2685
Db 721 TyHisAspArgAspAlaLeuGlyGlnGlyTyArgTyIleSerGlyGlyTySerLeu 740
2686 GGAGCAAACTCTTACTTTTGGATCATCGATTTTGGTCTAGCATTTACCGAAGTATTGGT 2745
Db 741 GlyAlaAsnSerTyPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760
2746 AGATCTAAAGATTATGTAGTGTCTCTTCCAAATCATCATCTGTCATAGGATCCGTTTAT 2805
Db 761 ArgSerLysAspTyValValCyAspSerAsnHisHisAlaCysIleGlySerValTy 780
2806 CTATCTACCCCAACAGCTTTTATGCGATCTTATTTGTTGCGAGATGCTTTATCCGTGCT 2865

Db 781 LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaIleargAla 800
Qy 2866 AGCTACGGGTTGGGAATCAGCATATGAAACCTCATATACATTTGCAGAGGAGCGAT 2925
Db 801 SerTyrGlyPheGlyAsnGlnHisMetIlySerTyrThrPheAlaGluGluSerAsp 820
Qy 2926 GTTCGTTGGGATAATACTGTCTGGCTGGAGAGATTGGAGGGGATTAACCAATTGTGATT 2985
Db 821 ValArgTTPAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
Qy 2986 ACTCCATTAAGCTTATTGATGAGTGGCTCTTCTGTCAGAGCTGAGTCTTCTTAT 3045
Db 841 ThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyr 860
Qy 3046 GCCGATCATGAATCTTTTACAGAGAGCGCATCAAGCTCGGGCATTCAGAGCGGACAT 3105
Db 861 AlaPheHisGluSerPheThrGluGlyAspGlnAlaArgAlaPheLysSerGlyHis 880
Qy 3106 CTCTTAATATCATAGTTCCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCT 3165
Db 881 LeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisPro 900
Qy 3166 AATAAATAGCTTTATGGCGCTTATCTGTGATGCTTATCGCACCATCTCTGTTACT 3225
Db 901 AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920
Qy 3226 GAGCAACGCTCTCATCCATCAAGAGACATCGACACAGATGCTCTTCATTAGCAAGA 3285
Db 921 GluThrThrLeuLeuSerHisGlnGluThrThrThrAspAlaPheHisLeuAlaArg 940
Qy 3286 CATCGAGTTGGTTAGAGATCTATGATGCTTCTTAACAAGTAATATAGAATATAT 3345
Db 941 HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyr 960
Qy 3346 GCCCATGAGATATGATATCGATGCTTCTCGAGGCTATGTTGAGTGCAGGAAGT 3405
Db 961 GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
Qy 3406 AGAGTCCGGTTC 3417
Db 981 ArgValArgPhe 984

RESULT 15

US-10-931-779-43
; Sequence 43, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-43

Alignment Scores:
Pred. No.: 0 Length: 984
Score: 5116.00 Matches: 984
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.9% Indels: 0
DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-43 (1-984)

Qy 466 GAAATCATGTCTCTCAAGGAATTTACATGGGAGAGCTTAACTGTATCATTTCCCTAT 525
Db 1 GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
Qy 526 ACTGTATTAGAGATCCGAGTGGGACTACTGTTTTCTTCGAGGAGAGTAAACATTAAAA 585
Db 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
Qy 586 AATCTTGACAATTTCTATTGCACTTTGCTTTAAAGTTGTTTTGGGAACTTTATAGGAGT 645
Db 41 AsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60
Qy 646 TTTACTGTTTAGGAGAGGACACTCGTTGACTTTCGAGAACATACGACTTCTACAAT 705
Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
Qy 706 GGGCAGCTCTAAGTAAATAGCGCTGATGAGCTGTTTACTATTAGGGTTTTAAAGAA 765
Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100
Qy 766 TTATCTCTTTTCCAAATTTCAATTTACTTCCCGTACTCTGCTGCTGCAACGACTTAATAAG 825
Db 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
Qy 826 GGTAGCCAGACTCCGACGACCAACTATACACCGTCTAATGCTATCTTTATCTTAAACA 885
Db 121 GlySerGlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThr 140
Qy 886 GATCTTTGTTACTCAATAAGAGAGTTCTCATTCTATAGTAAATTTAGTCTCTGAGAT 945
Db 141 AspLeuLeuLeuLeuAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160
Qy 946 GGGGAGCTATAGATGCTAAGAGCTTAAAGGTTCAAGGAATTAGCAAGCTTTGTCTTTC 1005
Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
Qy 1006 CAAGAAATATCTGCTCAAGCTGATGGGAGAGCTTCAAGTAGTACACAGTTTCTCTGT 1065
Db 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAla 200
Qy 1066 ATGGCTAACGAGGCTCTATTGCTTCTAGCGAATGTTGAGGAGTAGTAAGGGGAGGG 1125
Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
Qy 1126 ATTGCTGCTGTTTCCAGATGGGAGGAGGTGTCATCATCTACTTCAACAGAAATCCA 1185
Db 221 IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAspPro 240
Qy 1186 GTAGTAAGTTTTTCCAGAAATATCTCGGTAGAGTTTGATGGGAACGTAGCCGAGTAGA 1245
Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
Qy 1246 GGAGGGATTTACTCTCTAGGGAACTGCTTCTGTAATATGGAACCTTGTCTTCTC 1305
Db 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
Qy 1306 AACAAATGTTGCTTCTCTGTTTACATGCTGCTTAAGCAACCAACAGTGGACAGGCTTCT 1365
Db 281 AsnAsnValAlaSerProValTyrIleAlaLysGlnProThrSerGlyGlnAlaSer 300
Qy 1366 AATACGAGTAATAATTACGAGATGGAGAGCTATCTTCTGTAAAGATGGTGGCAGCA 1425
Db 301 AsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320
Qy 1426 GGATCCAACTCTGGATTCAGTTTCTTGTATGAGAGGGAGTAGTTTCTTTTAGTAGC 1485
Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSer 340
Qy 1486 AATGTAGCTGCTGGGAAAGGGGAGCTATTTATGCAAAAAGCTCTCGGTTGCTAACTGT 1545
Db 341 AsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLeuSerValAlaAsnCys 360

1546 GGCCTGTACAAATTTTAAAGGAATATCGTAAATGCTGGAGCGATTATTTAGGAGAA 1605
Db
361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380
1606 TCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATATTTTCGATGGGAATCTTAAA 1665
Db
381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400
1666 AGAACGCGCAAGAGAAGTCTGCCGATCTTAATGCGCTAACTGTCTCTCACAGCCATT 1725
Db
401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420
1726 TCGATGGGATCGGAGCGGAAATACGACATTAAGAGCTAAAGCAGCGCATCAGATTCTC 1785
Db
421 SerMetGlySerGlyGlyIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
1786 TTTAATGATCCCATCGAGATGGCAACCGAAATAACGACCAAGCCAGCGATCTTCCAACTT 1845
Db
441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
1846 CTAAAAATTAACGATGGTGAAGGATACACAGGGGATATGTTTTGCTAATGGAGCAGT 1905
Db
461 LeuLysIleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480
1906 ACTTTGTACCAAAATGTTACGATAGCAGCAAGGAGATTGTTCTTCGTGMAAAGCAAAA 1965
Db
481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
1966 TTATCAGTGAATTTCTTAAGTCAGACAGTGGGAGTCTGTATATGAAAGCTGGGAGTACA 2025
Db
501 LeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThr 520
2026 TGGGATTTGTAACTCCCAACACCAACAGACCTCTCGCGCTAAATCAGTTGATCAGC 2085
Db
521 TrpAspPheValThrProGlnProGlnGlnProAlaAlaAsnGlnLeuIleThr 540
2086 CTTTCCAAATCTCATTTGCTCTTCTTCTTTGTTAGCAAAACAATGCAGTTACGAATCCT 2145
Db
541 LeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThrAsnPro 560
2146 CCTACCAATCTCCAGCGCAAGATTCATCTCTGCAATCGTGGTAGCAAACTCTGCT 2205
Db
561 ProThrAsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGly 580
2206 TCTGTTACAAATAGTGGGCTATCTTTTGGAGATTTGGATGATACAGCTTATGATAGG 2265
Db
581 SerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArg 600
2266 TATGATTCGGTAGGTTCTAAATCAAAAAATCAATGCTCGAAATTAACAGTTAGGACTAAG 2325
Db
601 TyrAspTrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLys 620
2326 CCCCCAGCTTAATGCCCCATCAGATTTGACTCTAGGGAATGAGATGCTTAAGTATGCTAT 2385
Db
621 ProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyr 640
2386 CAAGGAAGCTGAAGCTTCGCTGGGATCCTAATACAGCAAAATAATGGTCTCTTATATCTCG 2445
Db
641 GlnGlySerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeu 660
2446 AAAGCTACATGACCTAAACTGGGTATATCTGGGCTCGAGCGTAGCTCTTTGGTT 2505
Db
661 LysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuVal 680
2506 CCAAAATAGTTTATGGGATCCCAATTTAGATATACGATCTGCGCATTCAGCAATTCAGACA 2565
Db
681 ProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700
2566 AGTGTGATGGGCGCTCTTATTTGTCGAGGATTAATGGGTTCTGGAGTTTCGAAATTTCTTC 2625
Db
701 SerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePhe 720
2626 TATCATGACCGCATCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATTCCTTA 2685

721 TyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeu 740
2686 GGACCAACTCTACTTTGGATCATCGATGTTTGTCTTAGCATTTACCGAAGTATTTGGT 2745
Db
741 GlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760
2746 AGATCTAAAGATTATGATGTGTCTGTTCCAAATCATCATCTTCATAGGATCCGTTTAT 2805
Db
761 ArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyr 780
2806 CTATCTACCCAAAGCTTTATGTTGGATCCCTATTGTTTCGGAGATGCGTTTATCCGTCT 2865
Db
781 LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAla 800
2866 AGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGAGAGAGCCAT 2925
Db
801 SerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAsp 820
2926 GTTCGTTGGGATAATAACTGTCTGCTGGAGAGATTGGAGCGGATTTACCGATTGTGATT 2985
Db
821 ValArgTrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
2986 ACTCCATCTAAGCTCTATTGAAAGAGTTGCGTCCCTTTCTGTCGCAAGCTGAGTTTCTTAT 3045
Db
841 ThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyr 860
3046 CCGCATCATGATCTTTTACAGAGCAAGCGCATCAAGCTCGGGCATTCGAAGCGGACAT 3105
Db
861 AlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHis 880
3106 CTCCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTTGATCGATGTTGTAGTACACATCT 3165
Db
881 LeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisPro 900
3166 AATAAATATAGCTTTATCGGGCTTATATCTGTGTATGCTTATCGCACCATCTCTGGTACT 3225
Db
901 AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920
3226 GAGCAACGCTCTATCCCATCMAGACATGGAACAACAGATGCTTCATTTTACCAAGA 3285
Db
921 GluThrThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArg 940
3286 CATGAGTGTGGTTAGAGGATCTATGATGCTTCTCTAAACAAGTAATATAGAAGTATAT 3345
Db
941 HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyr 960
3346 GSCCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGGTTTGAGTCAGGAAGT 3405
Db
961 GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
3406 AGAGTCCGGTTC 3417
Db
981 ArgValArgPhe 984
RESULT 16
US-10-498-327-17
; Sequence 17, Application US/10498327
; Publication NO. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 17
; LENGTH: 1013
; TYPE: PRT

ORGANISM: Chlamydia trachomatis
US-10-498-327-17

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5112.50 Matches: 977
Percent Similarity: 98.7% Conservative: 23
Best Local Similarity: 96.4% Mismatches: 12
Query Match: 64.9% Indels: 1
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-498-327-17 (1-1013)

QY	382	ATGCAACGCTTTCCCAATAGTTCTTTCTTCAATGATTTAGCTATTCTTGCTGCTCT	441
DB	1	MetGlnThrSerPheHsLysPhePheLeuSerMetLeuLeuAlaTyrSerCysCysSer	20
QY	442	TTAAATGGGGGGATATGACGACGAGAAATCATGCTTCTCAAGAAATTTACGATGGGAG	501
DB	21	LeuSerGlyGlyGlyTyrAlaAlaGluMetIleProGlnGlyIleTyrAspGlyGlu	40
QY	502	ACGTTAACTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTGTTT	561
DB	41	ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe	60
QY	562	TCTGACGAGAGTTAAACATTAATAAATCTTCACAAATCTTATGACGCTTTGCCCTTAAAT	621
DB	61	SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerLeuAlaLeuProLeuSer	80
QY	622	TGTTTTGGGAATTTATAGGGAGTTTACTGTTTTAGGAGGACACTCGTTGACTTTC	681
DB	81	CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe	100
QY	682	GAGAACATACGACTTCTACAAATGGGGCAGCTCTAGTAATAGCGTGTGATGGAGCTG	741
DB	101	GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu	120
QY	742	TTTACTATTGAGGTTTTAAAGATTATCTTTTCCAAATTCATTTACTTTGCCGTA	801
DB	121	PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal	140
QY	802	CTGCTGCTGCAACGACTTAATAAGGTTAGCCAGACTCCGACGACAACTATCACCGTCT	861
DB	141	LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrSerThrProSer	160
QY	862	AATGTTACTATTATTTCAACAGACTCTTTGTTTACTCAATATGAGAGTTCTCATTC	921
DB	161	AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe	180
QY	922	TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA	981
DB	181	TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln	200
QY	982	GGAATTAGCAGCTTTGTGTTCTTCAAGAAATACTGCTCAAGCTGATGGGGAGCTGT	1041
DB	201	GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys	220
QY	1042	CAAGTAGTCAACAGTTTCTGCTATGCTTAACGAGGCTCTATTGCCCTTTAGCGAAT	1101
DB	221	GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn	240
QY	1102	CTTGACGAGTAAGGGGGAGGATGCTGCTTTCAGATGGGACGAGGAGTGCTCA	1161
DB	241	ValAlaGlyValArgGlyGlyLeuAlaValGlnAspGlyGlnGlyValSer	260
QY	1162	TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTTCCAGAAATACTGCCGTAGAGTTT	1221
DB	261	SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe	280
QY	1222	GATGGGAACGTAGCCGAGTAGAGGAGGATTTACTCTACGAGGACGCTGCTCTCTG	1281
DB	281	AspGlyAsnValAlaArgValGlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu	300

QY	1282	ATAATGGAAAAACCTTGTTTCTCAACAATGTTGCTTCTCCTGTTTACATTGCTGCTAAG	1341
DB	301	AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrLeuAlaAlaGlu	320
QY	1342	CAACCAACAGTCGACAGGCTTCTAATACGAGTAATAATTTACGAGATGAGGAGCTATC	1401
DB	321	GlnProThrAsnGlyGlnAlaSerAsnThrSerAsnTyrGlyAspGlyGlyAlaIle	340
QY	1402	TTCTGTGAAGAAATGTCGCAAA---GCAGGATCCAATAACTCTCGATCAGTTCTCTTGAT	1458
DB	341	PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp	360
QY	1459	GGAGAGGAGTAGTTTCTTTTAGCAATGTAGCTGCTCGGAAAGGGGAGCTATTAT	1518
DB	361	GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr	380
QY	1519	GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGCAATATCGCTAAT	1578
DB	381	AlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn	400
QY	1579	GATGGTCGACGATTTTATTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATATGGA	1638
DB	401	AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly	420
QY	1639	GATATTATTTTGAATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAAT	1698
DB	421	AspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn	440
QY	1699	GGGTAACTGTGCTCACAAGCCATTTCCGATGGATCGGGGAGGAAATAACACACATTA	1758
DB	441	GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyPheLysIleThrThrLeu	460
QY	1759	AGACTAAAGCAGGCGCATCAGATTCTTTAATATCCCATCCGAGATGGCAACCGAAAT	1818
DB	461	ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn	480
QY	1819	AACGACGACGCGAGTCTTCCAACTTCTAAAAATTAAACGATGCTGAAGGATACACAGGG	1878
DB	481	AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyrThrGly	500
QY	1879	GATATTGTTTGTCTAATGGAAGCAGTACTTTGACCAAAATGTTTACGATAGACAGGA	1938
DB	501	AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly	520
QY	1939	AGGATTGTTCTTCGTGAAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCACAGGTGG	1998
DB	521	ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly	540
QY	1999	AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCAACAG	2058
DB	541	SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln	560
QY	2059	CCTCTCGGCTAATCAGTTCATCGCTTCCAACTGCAATGCTGCTTCTTCTTCTTCTTG	2118
DB	561	ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu	580
QY	2119	TTAGCAACAATGCAAGTTACGAATCTCTACCAATCTCCAGCGCAAGATTTCTCATCT	2178
DB	581	LeuAlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro	600
QY	2179	GCAGTCATTGTTAGCACAACTGCTGGTTCTGTTTACAAATTTAGTGGGCTATCTTTTTCG	2238
DB	601	AlaIleIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu	620
QY	2239	GATTTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTTCCTTAATCAAAAATCAAT	2298
DB	621	AspLeuAspAspThrAlaTyrAspArgTyrAspIlePheLeuGlySerAsnGlnLysIleAsp	640
QY	2299	GTCTGAAATACAGTTAGGGACTAAGCCGCCAGCTTAATGCCCATCAGATTGACTCTA	2358
DB	641	ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu	660
QY	2359	GGGAATAGATGCCCTTAAGTATGGCTATCAGGAGCTGGAAGCTGCGTGGATCCTAAT	2418

```

; Sequence 15, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/512,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-15

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.7% Mismatches: 15
Query Match: 64.8% Indels: 1
DB: 4 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-701-844-15 (1-1013)

QY 382 ATGCAACCTCTTCCATAAGTTCTTCTTCAATGATTCTAGCTTATTCTGCTCTCT 441
DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaLysCysCysSer 20
QY 442 TTAATGGGGGGGATATCCACAGAAATCATGGTTCCTCAAGAAATTTACGATGGGAG 501
DB 21 LeuAsnGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 AGTTAACTGTATCTTCCCTATCTACTGTATAGGAGATCCGAGTCCGAGTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGAGTTAACTTAAATAATCTTGCAATCTTATTCAGCTTTTGCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
QY 622 TGTTTTGGAACTTATTAGGAGTTTTTACTGTGTTTTAGGAGAGGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGAGCTTCTACNAATGGGCGAGCTCTAAGTAATAGCCGCTGCTGATGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCAATTTGCAATTCATTACTTGCCTGA 801
DB 121 PheThrIleGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140
QY 802 CTCGCTGCTGCACAGACTAATAGGGTAGCCAGACTCCGAGCTCCAGACACACATCTACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTTAAACAGATCTTTTCTTACTCAATATGAGAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTACTCTCTGGAGATGGGGAGCTATAGATCTAAGAGCTTAAACGGTTCAA 981
DB 181 TyrSerAsnSerValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTGTCTTCCAGAAATAATCTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 1000 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 3417
QY 3379 CGAGGCTATCGTTTGAAGTCAGGAGTAGAGTCGGTTC 3417
DB 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 17
US-10-701-844-15
```

Db 201 GlyIleSerLysLeuCyseValPheGlnGlnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTACACGAGTTCTCTGCTATGCTAACGAGGCTCTATTGCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheValAlaAen 240
Qy 1102 GTTCGAGAGTAGAGGGAGGAGTCTCTGCTCTAGGATGGCAGCAGGAGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaValGlnAenAspGlyGlnGlnGlyValSer 260
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
Qy 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCAGGAAACGTTGCTTCTCTG 1281
Db 281 AspGlyAenValAlaArgValGlyGlyGlyIleTySerTyGlyAenValAlaPheLeu 300
Qy 1282 AATAATGGAACCACTTCTGCTCAACAATGTGTGCTTCTCTGTTTATCATGCTGCTAAG 1341
Db 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyIleAlaAlaGlu 320
Qy 1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAenAenThrSerAenAenThrSerAenAen 340
Qy 1402 TTCTGTAGAAATGGTGCACAA---GCAGGATCCAAATACTCTGGATCAGTTTCTTTGAT 1458
Db 341 PheCysLysAenGlyAlaGlnAlaAlaGlySerAenAenSerGlySerValSerPheAep 360
Qy 1459 CGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGCTGTGGAAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTy 380
Qy 1519 GCCAAAAGCTCTCGTTGCTAACTGTGCGCTGTACAAATTTTAAAGAAATATCCCTAAT 1578
Db 381 AlAlaLysLysLeuSerValAlaAenCysGlyProValGlnLeuLeuGlyAenIleAlaAen 400
Qy 1579 GATGCTGAGCGCAATTTATTAGGAGATCTGGAGAGCTCAGTTTATCTGCTGATTTATGGA 1638
Db 401 AspGlyGlyAlaIleTyLeuGlyLeuSerGlyGlyLeuSerLeuSerAlaAepTyGly 420
Qy 1639 GATATTATTTTCAGTGGGATCTTAAAGAACAGCAAGAGATGCTGCCGATGTTAAT 1698
Db 421 AspMetIlePheAepGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAepValAen 440
Qy 1699 GCGTAACCTGTCTCTCAAGCCATTTCTGATGGATCGGAGGAGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyValIleThrThrLeu 460
Qy 1759 AGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGCGCAACGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAenAenProIleGluMetAlaAenGlyAen 480
Qy 1819 AACGAGCGAGCGAGTCTTCCAAACTTCAAAAATTAACGATGGTGAAGGATACACAGG 1878
Db 481 AenGlnProAlaGlnSerSerGluProLeuLysIleAenAepGlyGlyGlyTyThrGly 500
Qy 1879 GATATTGTTTTGCTTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAAGCA 1938
Db 501 AspIleValPheAlaAenGlyAenSerThrLeuTyGlnAenValThrIleGlnGly 520
Qy 1939 AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGly 540
Qy 1999 AGTCTGTATATGAGAGCTGGAGTACATGGGATTTTGTAACTCCCAACCAACCAACAG 2058
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAepPheValThrProGlnProGlnGln 560
Qy 2059 CTTCTGCGCGCTAATCAGTTGATCAGCTTTCCAATCTGCATTTGTCTCTTCTCTTTG 2118
Db 561 ProProAlaAlaAenGlnSerIleThrLeuSerAenLeuHisLeuSerLeuSerLeu 580

Qy 2119 TTACCAACAATGAGTTCAGATTCCGATTCCTACCAATTCCTCCAGCGCAAGATTTCTCATCT 2178
Db 581 LeuAlaAenAenAlaValThrAenProThrAenProThrAenProAlaGlnAenSerHisPro 600
Qy 2179 GCAGTCAATTGCTAGCACAACTGCTGCTCTCAATTAAGTGGGCTCTATCTTTTTCAG 2238
Db 601 AlaValIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
Qy 2239 GATTGGATGATCAGCTTATGATAGGTATGATTGGCTAGGTTCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAepAspThrAlaTyAspArgTyAspTrpLeuGlySerAenGlnLysIleAep 640
Qy 2299 GTCTGAAAATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAenAlaProSerAspLeuThrLeu 660
Qy 2359 GGGATGAGATGCTTAAGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGATCCTAAT 2418
Db 661 GlyAenGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAen 680
Qy 2419 ACAGCAATAATGCTCTTATCTCTGAAAGCTACATGAGCTAAACCTGGGTATAATCCT 2478
Db 681 ThrAlaAenAenGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAenPro 700
Qy 2479 GGGCTGAGCAGTAGTCTTCTTGGTCCAAATAGTTTATGGGATCCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAepIle 720
Qy 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGATGGCGCTCTTATTTGCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeu 740
Qy 2599 TGGGTTTCTGAGGATTCGAATTTCTCTCATCATGACCGCGATGCTTTTAGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAenPheTyIleAspArgAspAlaLeuGlyGlnGlyTy 760
Qy 2659 CGGTATATTAGTGGGGTATTCTTAGGAGCAAACTCTACTTTGGATCATCGATGTTT 2718
Db 761 ArgTyIleSerGlyGlyTySerLeuGlyAlaAenSerTyPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTAATGATGTGCTGTTCCAAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyValValCysArgSerAen 800
Qy 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTACCCAAACGCTTTATGTGATPCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyLeuSerThrLysGlnAlaLeuCysGlySerTy 820
Qy 2839 TTGTTTCGAGATGCGTTTATCCGTCTAGCTACGGGTTTGGGAATCAGCATATGAAAACC 2898
Db 821 ValPheGlyAepAlaPheIleArgAlaSerTyGlyPheGlyAenGlnHisMetLysThr 840
Qy 2899 TCATATACATTTGCAGAGCAGAGCGATGCTCGTTGGGATAATACTGTCTGCTCGAGAG 2958
Db 841 SerTyThrPheAlaGluGluSerAspValCysTrpAspAenAenCysLeuValGlyGlu 860
Qy 2959 ATTGAGCGGGAATTACCGAATTTGATTAATCTCATCTAAGCTCTATTGTAATGAGTTGCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyLeuAenGluLeuArg 880
Qy 3019 CTTTCTGCGAAGCTGAGTTTCTTATGCGGATCATGAAATCTTTTACAGAGAAAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyAlaAepHisGluSerPheThrGluGluGlyAep 900
Qy 3079 CAAGCTCCGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCGATGTTCTAGTACATCCTTAATAATAGCTTTTATCGCGGCTTATATCTGT 3198
Db 921 PheAepAArgCysSerSerThrHisProAenLysTySerPheMetGlyAlaTyIleCys 940


```
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyThrGly 500
Qy 1879 GATATTGTTTGGTAAGCAGTACTTTGTACCAAAATGTTACATAGACAGCA 1938
Db 501 AspileValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
Qy 1939 AGGATTGTTCTCGTGAAGCAAAATATCATAGTGAATTCCTAAGTCACAGAGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
Qy 1999 AGTCGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACCAACCAACCAAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
Qy 2059 CCTCTGCCCTAAATCAGTTGATCAGCTTTCCAACTGCAATTTGTCTCTTCTTTG 2118
Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
Qy 2119 TTAGCAAAATGCGATTACGAATCTCTTACCAATCTCTCAGCGCAAGATCTCATCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProThrAsnProAlaGlnAspSerHisPro 600
Qy 2179 GCAGTCATTGGTAGCACAACTGCTGGTCTGTTCATCAATTAGTGGGCCCTATCTTTTCAG 2238
Db 601 AlaValIleGlySerThrThrAlaGlyProValThrIleSerGlyProPhePheGlu 620
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATGGCTAGGTCTTCAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
Qy 2299 GTCCTGAAATTACAGTTAGGACTAAGCCCGCAGCTAATGCCCATCAGATTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
Qy 2359 GGGATGAGATGCCTAAGTATGGCTATCAAGGAAGCTGGAAGCTGGTGGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyThrLeuLysAlaThrTrpLysThrGlyTyrAsnPro 680
Qy 2419 ACAGCAAAATATGGTCTTATCTGAAAGCTACATGAGCTAAAGCTGGGTATATCCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpLysThrGlyTyrAsnPro 700
Qy 2479 GGGCTGAGCGAGTACTTCTTGGTCCAAATAGTTTATGGGGATTCATTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspile 720
Qy 2539 CGATCTCGCATTCAGCAATTCAGCAAGTGGATGGCGCTCTTATTGTGAGGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy 2599 TGGGTTTCTGGAGTTTCGAATTTCTTATCATGACCGCGATGCTTTAGTCAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPheSerTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy 2659 CGGTATATTAGTGGGGTTATTCCTTAGGAGCAAACTCTTACTTTGGATCATCGATTT 2718
Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCGAGTATTTCGTAGATCTAAAGATTATGTAGTGTGCTTCCCAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy 2779 CATCATGCTGCATAGGATCCGTTTATCTATCTATCCCAACAGCTTTATGTGGATCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTCCGGAGATCGTTTATCCGTCTAGCTAGCGTTTGGGAATCAGCATATGAACCC 2898
Db 821 LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2899 TCATATACATTTGCAGAGGAGCATGTTCTGGGATAATTAACCTGCTGCGTCGAGAG 2958
```

```
Db 841 SerTyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGGAGGGGATTACCGATTGTGATTAATCTCCATCTAAGCTCTATTGGAATGAGTTGGT 3018
Db 861 IleGlyValGlyLeuProIleValThrThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy 3019 CCTTTCTGTCAGCTGAGTTTCTTATGCCGATCATGAATCTTTTACAGAGGAAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGCATTCGAAGAGCGGACATCTCCATAATCTATCAGTTCTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTGATGATGATCTTCTAGTACATCCTTAATAATATAGCTTTATGCGCGCTTATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCCACCATCTCTGCTAGTACAGCAAGCTCCCTATCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTTGTGTTAGAGGATCTATGTATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy 3319 TCTCTAACAAATATATAGAAATATATGCGCATCGAAGATATGATATGAGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGCTGATGTTTGGTGGCAGGAAGTAGATCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 19
US-10-766-711-15
; Sequence 15, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-15

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.7% Mismatches: 15
Query Match: 64.8% Indels: 1
Db: 4 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-766-711-15 (1-1013)
Qy 382 ATGCAAAAGCTCTTTCCATAAGTCTTTCTTTCAATGATTTAGCTTATTTCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
Qy 442 TTAATATGGGGGGGATATGCGACGAATATCATGTTCTCTCAAGGAATTTACATGGGGAG 501
Db 21 LeuAsnGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
```

QY 502 ACGTTAACTGTATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGACTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGAGGAGAGTTAACTTAAATAAATCTTGCAAAATCTTATTGCGAGCTTTGCTTTAAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAACCTTATTAGGAGATTTACTGTTTATAGGAGGAGACACTCGTTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGACTTCTCAAAATGGGCGAGCTCTAAGTAATAGCGCTCTCATGACTG 741
DB 101 GluAsnIleArgThrSerThrAsnGlyAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAATTATCTTTTCCAATTGCAATTCATTCTGCGGTA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140
QY 802 CTGCGCTGCGAACGACTAATAAGGTAGCCAGACTCCGACGACACATCTACACCGTCT 861
DB 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrProSer 160
QY 862 AATGGTACTATTATTCTAAACACAGATCTTTTGTACTCAATAATGAGAAGTTCTCATTC 921
DB 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnGlnLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA 981
DB 181 TyrSerAsnSerValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTCTTCCAAAGAAATATCTGCTCAAGCTGATGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTCACCAAGTTCTCTGCTATGGCTTAACGAGCTCTTATTCCTTTGAGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTCAGGAGTAGAAGGGGAGGAGTGTCTGTCTGTTTCCAGATGGGAGGAGGTGTCA 1161
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAAACAGAGATCCAGTAGTAAGTTTTCGAAATATCGCGTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGNACGTAGCCCGAGTAGGAGGAGGATTTACTCTACCGGAGCTGTCTTCCCTG 1281
DB 281 AspGlyAsnValAlaAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAAAAACCTTTGTTCTCAAAATGTTGCTTCTCTCTGTTTACATTTGCTGTAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320
QY 1342 CAACCAACAAAGTGGACAGGCTTCTTAATACAGTAATAATTAACGAGATGAGGAGCTATC 1401
DB 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTTCTGTAAGATGGTGGCGAA---GCAGGATCCATAACTCTGGATCAGTTTCTTTTGTAT 1458
DB 341 PheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGAGGAGTAGTTTCTTTTATGAGCAATGTAGCTCTGCGAAAGGGGAGCTATTAT 1518
DB 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyr 380
QY 1519 GCCAAAAGCTCTCGGTGCTTAACCTGTGGCCCTGTACAAATTTTTTAAGGAATATCGCTAAT 1578
DB 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnLeuLeuGlyAsnIleAlaAsn 400
QY 1579 GATGGTGGAGCGATTATTTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGA 1638

DB 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTATTTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAATGCTGCCGATGTTAAT 1698
DB 421 AspMetIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGCTAACTGTGCTCTCAACGCCATTTTCGATGGGATCGGAGGAGAAATAACAGCATTA 1758
DB 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGCGCAACCGAAT 1818
DB 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACGAGCCAGCGCAGCTCTTCCAAACTTCTAAAAATTAACGATGGTGAAGGATACACAGGG 1878
DB 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleLeuAsnAspGlyGluGlyTyrThrGly 500
QY 1879 GATATTGTTTTTGTCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGAGCAAGA 1938
DB 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTGTTCTTCGTGAAAAGGCCAAAAATTATCAGTGAATTTCTTAAGTCAGACAGTGG 1998
DB 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGAATCTTAACTCCACACACCAACAG 2058
DB 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CTTCTGCGCGCTTAATCAGTTGATCAGCTTTTCCAATCTGCATTTGTCTCTTCTTCTTTG 2118
DB 561 PropAlaAlaAsnGlnSerIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
QY 2119 TTAGCAAAACAATGCAAGTACGAATCTCTACAAATCTCCAGCGCAAGATTTCTCATCCT 2178
DB 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCATTGTTAGCACACTCTGCTGTTCTGTTTACAATTAGTGGGCTATCTTTTGTAG 2238
DB 601 AlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTTGGAATGATACAGCTTATGATAGTATGATGCTAGTTTCTAATCAAAAATCAAT 2298
DB 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCCTGAAATTAAGTTAGGAGCTAAGCCCGCAGCTAATGCCCCCATCAGATTTGACTCTA 2358
DB 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGAAATGAGATGCTTAAGTATGGCTATCAGGAGCTGGAAGCTTCGCTGGGATCCTAAT 2418
DB 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680
QY 2419 ACAGCAAAATAAGTCTCTTATCTCTGAAAGCTACATGACTAAAACTGGGTATAATCCT 2478
DB 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCCTGAGCGAGTAGCTTCTTTGTTTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
DB 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720
QY 2539 CGATCTGCGGATTCAGCAATTCAGCAGTGTGATGGGCGCTCTTATTTCTCGGAGTTA 2598
DB 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTCTATCATGACGCGATGCTTTTAGTGTACAGGATAT 2658
DB 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
QY 2659 CGGTATATTAGTGGGGTTATTCTCTTAGGAGCAAACTCTTACTTTGGATCATCGATGTTT 2718

Db 761 ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCAAGTATTTGGTAGACTAAAGATATGATAGTGTGCTTCCAAT 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy 2779 CATCATGCTTCATAGGATCCGTTTATCTATCTATCCCAACCAAGCTTTATGGATCCCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTCCGAGATCGCTTTATCCGCTAGCTACGGGTTTGGCAATCCAGCATATGAAACC 2898
Db 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2899 TCATATACATTTGCAGAGAGCGATGTTCCGTTGGGATAATACTGCTGGCTGGAGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValCysTyrAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGGAGCGGATTTACCGATTGTGATTACTCCATCTAAGCTTATTTGAATGAGTTGCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGlnLeuArg 880
Qy 3019 CCTTCGTCGAGCTGAGTTTCTTATGCGATCATGAATCTTTTACAGAGGAGCGAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATGTTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTGTATCATGTTCTAGTACATCATCTTAATAATATAGCTTTATGCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCGCACCATCTCTGTAATGAGACACGCTCTATCCATCCATCAAGAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrItp 960
Qy 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGGAGTGTGGTGTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
Qy 3319 TCTCTAACAGTAATAATAGAGTATATGCGCATGGAAGATATGATGATGATGCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGCTATGTTTCAGTGCAGAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 20

US-10-766-711-16
; Sequence 16, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRF
; ORGANISM: Chlamydia sp.
US-10-766-711-16

Alignment Scores:

Pred. No.: 0 Length: 1013

Score: 5106.50 Matches: 980
Percent Similarity: 98.2% Conservative: 15
Best Local Similarity: 96.7% Mismatches: 17
Query Match: 64.8% Indels: 1
DB: 4 Gaps: 1
US-10-701-844-1 (1-4435) x US-10-766-711-16 (1-1013)
Qy 382 ATGCAACAGCTTTTCCATAAGTCTTTCTTTCAATGATCTAGCTTATCTTGTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLeuPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
Qy 442 TTAATGGGGGGGATATGACAGCAAAATCATGCTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuThrGlyGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
Qy 502 ACGTTAACTGATCATTTCCCTATCTATCTATAGGAGATCCGAGTGGGACTACTCTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTGAGGAGAGTTAAACATTAATAAATCTTGACAAATCTTATGCAAGCTTTGCTTTAAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
Qy 622 TGTTTTGGAACTTATTAGGAGGTTTACTGTTTATAGGAGGAGGACACTGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATPACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTACTATGAGGGTTTAAAGAATTATCTTTTCCAAATTCCAATTCATTAATCTTCCCTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACCAACTCTACACCGCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrProSer 160
Qy 862 AATGTACTATTATTCTTAAACAGACTCTTTTGTGTACTCAATATATGAGAAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnGlnLysPheSerPhe 180
Qy 922 TATAGTAATATAGTCTCTGGAGAGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln 200
Qy 982 GGAATTAGCAAGCTTTGTCTCTCCAGAAATACTGCTCAAGCTGATGGGGAGCTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
Qy 1042 CAAGTAGTCACCAAGTTCTCTGCTATGCTTAAACAGAGCTCTTATGCTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsn 240
Qy 1102 GTTCAGAGTAAGAGGGGGAGGATTCCTGCTGTTCCAGGATGGCAGAGGAGTGTCA 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
Qy 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATACCTCCGCTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
Qy 1222 GATGGAAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGAAACGTTCTTCTTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
Qy 1282 AATAATGAAAAACCTTGTCTCAACAATGTGTCTTCTCTGTTTACATTCGCTGAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyrIleAlaAlaGlu 320
Qy 1342 CAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATTACGAGATGGAGAGCTATC 1401

Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnThrGlyAspGlyGlyAlaIle 340
 QY 1402 TTCTGTAAGAATGGTGGCGCAA---GCAGGATCAATACTCTGGATCAGTTTCCTTTGAT 1458
 Db 341 PheCysLeuAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
 QY 1459 GGAGGGAGTAGTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
 Db 361 GlyGluGlyValValPhePheSerAsnValAlaAlaGlyGlyGlyGlyAlaIleIleIyr 380
 QY 1519 GCCAAAAAGCTCTCGGTTGCTTAACCTGGCCCTGTACAAATTTTTTAAGGAATATCGCTAAT 1578
 Db 381 AlaLeuLeuLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsn 400
 QY 1579 GATGTGGAGCGAATTTATTTAGAGAAATCTGGAGAGCTCAGTTTATCTCTGATATGGA 1638
 Db 401 AspGlyGlyAlaIleIyrLeuGlyGlySerGlyGlyLeuSerLeuSerAlaAspIyrGly 420
 QY 1639 GATATTATTTTCGATGGGAATCTTAAAGCAACAGCCAAAGAGAAATGCTCCGATGTTAAT 1698
 Db 421 AspIleIlePheAspGlyAsnLeuIyrArgThrAlaIleGlyGlnAlaAspValAsn 440
 QY 1699 GCGTAACTGTGCTCACAGCCATTTCGATGGGATCGGAGGAGGAGAAATAACGACATTA 1758
 Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460
 QY 1759 AGAGCTAAAGCGGCGATCAGATTCTTTTAATGATCCCAATCGAGATGGCAACGGAAT 1818
 Db 461 ArgAlaLeuAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
 QY 1819 AACCAGCCAGCGAGTCTCCAAACTTCTAAATAATTAACGATGGTCAAGGATACACAGG 1878
 Db 481 AsnGlnProAlaGlnSerSerGluProLeuLeuIleAsnAspGlyGlyGlyIyrThrGly 500
 QY 1879 GATATTGTTTCTTAATGAAGCAGTACTTCTTACCAAAATCTTACGATAGAGCAAGGA 1938
 Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuIyrGlnAsnValThrIleGluGlnGly 520
 QY 1939 AGGATTGTTTCTCGTGAAGAGCAAAATTAATCAGTGAATTTCTTAAGTCAGAGAGTGGG 1998
 Db 521 ArgIleValLeuArgGluIyrAlaLeuLeuSerValAsnSerLeuSerGlnThrGlyGly 540
 QY 1999 AGTCTGTATGGAAGCTGGAGTACATGGGATTTGTAACTCTCCACCAACCAACACAG 2058
 Db 541 SerLeuIyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
 QY 2059 CCTCTGCGCTAATCAGTTGATCAGCTTTCCAACTCTGCAATTTGTCTCTTTCTTTCTT 2118
 Db 561 ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
 QY 2119 TTAGCAACCAATGAGTTACGAATCTCTCCACCAATCTCCAGCGCAAGATTTCTCATCT 2178
 Db 581 LeuAlaAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisPro 600
 QY 2179 GCAGTCATGCTAGCACACTGCTGTTCTGTACAAATTAAGTGGGCTATCTTTTGTAG 2238
 Db 601 AlaValIleGlySerThrThrAlaGlyProValThrIleSerGlyProPhePheGlu 620
 QY 2239 GATTTGGATCATCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298
 Db 621 AspLeuAspAspThrAlaIyrAspArgIyrAspIyrAspIyrAspIyrAspIyrAsp 640
 QY 2299 GTCTGAAATTAACAGTTAGGAGCTTAAGCCCCAGCTAATGCCCATCAGATTTGACTCTA 2358
 Db 641 ValLeuLeuLeuGlnLeuGlyThrGlnProSerAlaAlaAlaProSerAspLeuThrLeu 660
 QY 2359 GGGAAATCAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTGGTGGATCTCTAAT 2418
 Db 661 GlyAsnGluMetProIyrIyrGlyIyrGlnGlySerIyrIyrIyrIyrIyrIyrIyrIyr 680
 QY 2419 ACAGCAATTAATGCTCTTACTCTCAAGCTACATGAGTAACTGAGTAACTGATTAATCCT 2478

Db 681 ThrAlaAsnAsnGlyProIyrThrLeuLysAlaThrIyrThrIyrLeuThrGlyIyrAsnPro 700
 QY 2479 GGGCTCTGAGCAGTAGTACTTCTTTGGTCCAAATAGTTTATGGGATCCATTTTAGATATA 2538
 Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuIyrIyrGlySerIleLeuAspIle 720
 QY 2539 CGATCTGCGCATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 2598
 Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerIyrCysArgGlyLeu 740
 QY 2599 TGGGTTTCTGGAGTTTCGAATTTCTTCTCATCATGACCGCATGCTTTAGGTCAAGGATAT 2658
 Db 741 TrpValSerGlyValSerAsnPheSerIyrHisAspArgAspAlaLeuGlyGlnGlyIyr 760
 QY 2659 CGGTATATTAGTGGGGTATTCTTCTAGGAGCAAACTCTTCTTCTTCTTCTTCTTCTTCT 2718
 Db 761 ArgIyrIleSerGlyGlyIyrSerLeuGlyAlaAsnSerIyrPheGlySerSerMetPhe 780
 QY 2719 GGTCTAGCAATTCACGAAGTATTTGGTAGATCTTAAAGATATATGATGCTGCTTCTCAAT 2778
 Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspIyrValValCysArgSerAsn 800
 QY 2779 CATCATGCTGTAGTATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2838
 Db 801 HisHisAlaCysIleGlySerValIyrLeuSerThrLysGlnAlaLeuCysGlySerIyr 820
 QY 2839 TTGTTCCGAGAGTCGTTTATCCGTGCTAGCTACGGGTTTGGGAATCAGCATATGAAAACC 2898
 Db 821 LeuPheGlyAspAlaPheIleArgAlaSerIyrGlyPheGlyAsnGlnHisMetLysThr 840
 QY 2899 TCATATACATTTCCAGAGAGCGATGTTGCTGGGATAATACTGCTGCTGCTGAGAG 2958
 Db 841 SerIyrThrPheAlaGluGluSerAspValArgIyrAspAsnAsnCysLeuValGlyGlu 860
 QY 2959 ATTGGAGCGGATACCGATTGCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3018
 Db 861 IleGlyValGlyLeuProIleValThrThrProSerLysLeuIyrLeuAsnGluLeuArg 880
 QY 3019 CCTTTCGTCGACGCTCAGTTTCTTCTTCTGCGATCATGAATCTTTTACAGAGGAGCGAT 3078
 Db 881 PropheValGlnAlaGluPheSerIyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
 QY 3079 CAAGCTCGGCGATTCAGAGCGGACATCTCTTAAATCTATCAGTTCTCTGTTGAGTGAAG 3138
 Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
 QY 3139 TTTGATCGATGTTCTAGTACATCTCTTAAATATATAGCTTTTATGCGGCTTATATCTGT 3198
 Db 921 PheAspArgCysSerSerThrHisProAsnLysIyrSerPheMetGlyAlaIyrIleCys 940
 QY 3199 GATGCTTATCGCACCATCTCTGCTACTGAGACAACTCTCTATCCCATCAAGAGACATGG 3258
 Db 941 AspAlaIyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
 QY 3259 ACAACAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3318
 Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetIyrAla 980
 QY 3319 TCTCTAACAGTAAATATAGAGTATATGCGCATGGAAGATATGAGTATCGAGATGCTTCT 3378
 Db 981 SerLeuThrSerAsnIleGluValIyrGlyHisGlyArgIyrGlyIyrArgAspThrSer 1000
 QY 3379 CGAGGCTATGTTTGGTGCAGGAGTATGAGTCCGGTTC 3417
 Db 1001 ArgGlyIyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 21
 US-10-931-779-15
 ; Sequence 15, Application US/10931779
 ; Publication No. US20050048557A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jackson, W. James
 ; APPLICANT: Pace, John

```
/ TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
/ FILE REFERENCE: BP104
/ CURRENT APPLICATION NUMBER: US/10/931,779
/ CURRENT FILING DATE: 2004-09-01
/ PRIOR FILING DATE: 09/542,520
/ PRIOR FILING DATE: 200-04-03
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 15
/ LENGTH: 1013
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-10-931-779-15

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.7% Mismatches: 15
Query Match: 64.8% Indels: 1
DB: 5 Gaps: 1

US-10-701-844-1 (1-4435) x US-10-931-779-15 (1-1013)
QY 382 ATGCAAGCTCTTTCCATAAGTCTCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTySerCysSer 20
QY 442 TTAATGGGGGGATATCGACAGCAAAATCATGGTTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuAenGlyGlyTyAlaAlaGluIleMetValProGlnGlyIleTyAspGlyGlu 40
QY 502 ACGTTAACTGATCAATTTCCCTATACTGTATAGGAGATCCGAGTGGGACTACTGTTTT 561
Db 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGAGGAGGTTAACTTAAAAATCTTGACAAATCTTATGCAAGTTTGGCTTTAAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAACCTATTAGGGAGTTTACTGTTTTAGGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741
Db 101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTACTTCCGCTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTTAATAAGGTAGCCAGACTCCGACACACACATCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrThrProSer 160
QY 862 AATGGTACTATTATTCTAAAAACAGATCTTTGTTTACTCAATAATAGAGAGTTTCTCATTC 921
Db 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTATGCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA 981
Db 181 TyrSerAsnSerValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGCTCTTCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAspGlyAlaCys 220
QY 1042 CAAGTAGTCACCAAGTTTCTGCTATGGCTTAACGAGGCTCTTATTCCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTGCAGGAGTAAGAGGGGAGGATTCGCTGTTTCAGGATGGGCGAGGAGTGTCATCA 1161
```

```
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCACAGAAATACGCGTAGAGTTT 1221
Db 261 SerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGGAACGTTCTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAAAAACCTGTTTCTCAACAATGCTCTCTCTCTCTGTTTACATTGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyIleAlaAlaGlu 320
QY 1342 CAACCAACAGCTGGACAGGCTTCTAATACGAGTAAATAATACGAGATGGAGGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIle 340
QY 1402 TTCGTGAAGAATGGTGGCGCAA---GCAGGATCCCAATAACTCTGGATCAGTTCTCTTCAAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAsp 360
QY 1459 GGAGGAGGAGTAGTTTCTTTAGTACCAATGCTGCTGGGAAAGGGGAGCTATTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyValAlaIleTy 380
QY 1519 GCCAAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAsnCysGlyProValGlnLeuLeuGlyAsnIleAlaAsn 400
QY 1579 GATGGTGGAGCGATTTATTAGGAGATCTCGAGAGCTCAGTTTATCTCTGCTGATTATGA 1638
Db 401 AspGlyGlyAlaIleTyLeuGlyLeuSerGlyGluLeuSerLeuSerAlaAspTyGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCCGATGCTTAAT 1698
Db 421 AspMetIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GCGCTAACTGTGCTCTCAAGCCATTTTCGATGGGATCGGAGGGGAAATTAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGATCAGATTCCTTTTAAATGATCCATCGAGATGGCAACGGAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1818 AACCCAGCCAGCGAGTCTCCAAACTTCTAAAAATTAACGATGGTGAAGGATACACAGGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGlyTyThrGly 500
QY 1879 GATATTGTTTTCGCTAAATGGAAGCAGTACTTTGTACCAAAATGTTTACATAGACAGAGA 1938
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyGlnAsnValThrIleGluGlnGly 520
QY 1939 AGGATTGTTCTTCGTGAAGGCAAAATTAATCAGTGAATTCCTTAAGTCAGACAGGTTGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCTGTATATGGAAGCTGGGAGTACATGGGATTTTGTAACTCCACACACCACCAACAG 2058
Db 541 SerLeuTyMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
QY 2059 CCTCTGCGCGCTTAATCAGTTGATCAGCTTCTCAATCTCGATTTGCTCTTCTCTTTG 2118
Db 561 ProProAlaAlaAsnGlnSerIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAAACAATGAGTTACGATCTCTCTACCAATCTCTCCAGCGCAAGATTCATCTCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProThrAsnProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCATTGTTAGCACAACTGCTGTTCTGTACAAATCTAGTGGGCTATCTTTTGTAG 2238
```


Db 601 AlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePheGlu 620
Qy 2239 GATTGGATGATACAGCTTATGATAGGTATGATTCGCTAGGTCTTAATCAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnValAsp 640
Qy 2299 GTCCCTGAAATACAGTTAGGACTAAGCCCGAGCTAATGCCCATCAGATTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProAlaAsnAlaProSerAspLeuThrLeu 660
Qy 2359 GGGATGAGATCCCTAAGTATGGCTATCAAGGAGCTGGAAGCTGGCGGATCCTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
Qy 2419 ACAGCAATATAGTCTTATCTGTAAGAGCTACATGGACTAAACTGGGTATATCCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpLysThrGlyTyrAsnPro 700
Qy 2479 GGGCTGAGCGAGTAGTCTTTGGTCCAAATAGTTTATGGGATCCATTTTATAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
Qy 2539 CCATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTGTCGAGATTA 2598
Db 721 ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
Qy 2599 TGGGTTCTGGAGTTTCCGAATTTCTTCTATCATGACCGGATGCTTTAGTCTAGGATAT 2658
Db 741 TrpValSerGlyValSerAsnPheTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
Qy 2659 CGGTATATTAGTGGGGTATTCTCTTAGGAGCAACTCTTCTTCTGATCATCGATGTTT 2718
Db 761 ArgTyrIleSerGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
Qy 2719 GGTCTAGCATTTACCAAGTATTGTTGGTAGATCTAAAGATTATGTAGTGTGTTCCAA 2778
Db 781 GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800
Qy 2779 CATCATGTTGCATGAGTCCGTTTATCTATCTACCAACAGCTTTATGTGATCTCTAT 2838
Db 801 HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
Qy 2839 TTGTTGGAGATCGGTTTATCCGTGCTAGCTAGCTGGTGGGATCAGCATATGAAAC 2898
Db 821 ValPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
Qy 2899 TCATATACATTTCCAGAGAGAGCGATGTTGCTGGGATAATAACTGCTCTGGCTGGAG 2958
Db 841 SerTyrThrPheAlaGluSerAspValCysTrpAspAsnAsnCysLeuValGlyGlu 860
Qy 2959 ATTGGAGCGGGATTACCGATTGTGATTCTCCATCTAAGCTCTATTGTAAGTTGCGT 3018
Db 861 IleGlyValGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuArg 880
Qy 3019 CTTTCTGTCAGCTAGTTTCTTATGCCGATCATGAATCTTTTACAGAGAGCGCAT 3078
Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
Qy 3079 CAAGCTCGGCGCATTCAGAGCGGACATCTCTAAATCTATCATCTCTGCTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
Qy 3139 TTTGATCGATGTTCTAGTACACATCCCTAATAATATAGCTTTTATGGCGGCTTATATCTGT 3198
Db 921 PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
Qy 3199 GATGCTTATCGACCATCTCTGTTACTGACACACGCTCTCTATCCCATCAAGACATGG 3258
Db 941 AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
Qy 3259 ACAACAGATCCCTTACATTTAGCAAGATGGAGTCTGTTAGTATGAGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980

Qy 3319 TCTCTAACACAGTAAATATAGAAATATATGGCCATCGAAGATATGAGTATCGAGATCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
Qy 3379 CGAGCTATGTTTTCAGTGCAGGAGTATAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
RESULT 22
US-10-931-779-16
; Sequence 16, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-16
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.2% Conserved: 15
Best Local Similarity: 96.7% Mismatches: 17
Query Match: 64.8% Indels: 1
DB: 5 Gaps: 1
US-10-701-844-1 (1-4435) x US-10-931-779-16 (1-1013)
Qy 382 ATGCAAAACGCTTCCATTAAGTCTTCTTCTTCAATGATTCTAGCTTATTCTTCTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysSer 20
Qy 442 TTAATATGGGGGGATATGCAGCAAAATCATGTTCTCAAGGAATTTACGATGGGAG 501
Db 21 LeuThrGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
Qy 502 ACCTTAACGTATCATTTCCCTATCTATAGGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
Qy 562 TCTGCAGGAGATTAACTTAAATACTTGCAATCTTATTCACGCTTTCCTTAACT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
Qy 622 TGTCTTGGCACTTATAGGAGTCTTCTGTTTATAGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
Qy 682 GAGAACATACGACTTCTCAAAATGGCGAGCTCTAAGTAATAGCGCTGCTGATGGAATG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaLeuSerAspSerAlaAsnSerGlyLeu 120
Qy 742 TTTTCTATTGAGGGTTTTAAAGAAATTAATCTTTTCCAAATTTGCAATTTCACTTCCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
Qy 802 CTCCTGCTGCAACGATTAAGGTAGCCAGACTCCGACGACCAATCTACACCGTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrThrThrProSer 160
Qy 862 AATGGTACTATTATTCTAAACACAGATCTTTTGTACTCAATTAATGAGAAAGTTCTCATTC 921

QY 3079 CAAGCTGGGATTCACAGCGGACATCTCCTAAATCTATCAGTTCTCTGTTGGAGTGAAG 3138
Db GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValIys 920
QY 3139 TTTGATCGATGTTCTAGTACACATCCTTAATAATATAGCTTTATGGCGGCTTATATCTGT 3198
Db PheAspArgCysSerSerThrHisProAsnIysTyrSerPheMetGlyAlaTyrIleCys 940
QY 3199 GATGCTTATCCACCATCTCTGGTACTGAGACAACGCTCCTATCCATCAAGACATGG 3258
Db AspaIatyrArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTCTATTTAGCAAGACATGGAGTTGTTAGAGGATCTATGTATGCT 3318
Db ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValAlaGlySerMetTyrAla 980
QY 3319 TCTCTAAACAAGTAATATAGAGTATATGCGCATGGAAGATATGATATCGAGATGCTTCT 3378
Db SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGCTATGTTGAGTGCAGGAAGTAGATCGGTTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerIysValArgPhe 1013

RESULT 23

US-09-841-132-190
; Sequence 190, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yahir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: POSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-190

Alignment Scores:
Pred. No.: 0 Length: 1006
Score: 5090.00 Matches: 979
Percent Similarity: 99.9% Conservative: 2
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 64.6% Indels: 0
DB: 3 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-841-132-190 (1-1006)

QY 472 ATGGTTCTCAAGAAATTTACAGTGGGAGAGCTTAACGTGTATCATTTCCCTATACTGTT 531
Db 25 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44
QY 532 ATAGGATCCGAGTGGGACTACTGTTTTCTCGCAGGAGAGTTAACTTAAATAATCTT 591
Db 45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuIysAsnLeu 64
QY 592 GACAATTTCTATTCAGCTTTCCTTAAGTTGTTTGGGAACCTATTAGGAGATTTTACT 651
Db 65 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84
QY 652 GTTTTAGGAGGAGACATCTGTTGACTTTTCGAGAACATACGAGCTTCTCAAAATGGGCA 711
Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104
QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTGAGGGTTTTAAAGAAATATCC 771

Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheIysGluLeuSer 124
QY 772 TTTTCCAAATGCAATTCATTACTTGCCTGCTGCTCAACGACTAAATAAGGGTAGC 831
Db 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnIysGlySer 144
QY 832 CAGACTCCGACGACACATCTACACCGTCTAATGGTACTATTATTCTTAAACAGATCTT 891
Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerIysThrAspLeu 164
QY 892 TTGTACTCAATATGAGAGTTCTCATTTATAGTAAATTTAGTCTCTGGAGATGGGGA 951
Db 165 LeuLeuLeuAsnAsnGluIysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 184
QY 952 GCTATAGATGCTAAGAGCTTAACGCTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAAGAA 1011
Db 185 AlaIleAspAlaIysSerLeuThrValGlnGlyIleSerIysLeuCysValPheGlnGlu 204
QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTTGTCAAGTAGTCACAGTTTCTCTGTATGCT 1071
Db 205 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 224
QY 1072 AACGAGGCTCTATTGCTTGTAGCGAATGTTGCAGAGTAAAGGGGGAGGATTGCT 1131
Db 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 244
QY 1132 GCTGTTCCAGGATGGCAGCAGGAGTGTCAATCATCTACTTCAACAGAGATCCAGTAGTA 1191
Db 245 AlaValGlnAspGlyGlnGlnGlyValSerSerThrThrSerThrGluAspProValVal 264
QY 1192 AGTTTTCACAGAAATATCTGCGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGAGG 1251
Db 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 284
QY 1252 ATTTACTCCTACGGGAACGTTGCTTCTGTAATATGAAAAACCTTGTCTTCAACAT 1311
Db 285 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyIysThrLeuPheLeuAsnAsn 304
QY 1312 GTTGCTTCTCTGTTTACATTCGTGTAAGCAACCAACGAAGTGGACAGCTTCTAATACG 1371
Db 305 ValAlaSerProValTyrIleAlaAlaIysGlnProThrSerGlyGlnAlaSerAsnThr 324
QY 1372 AGTAATAATACGAGATGGAGGAGCTATCTTCTTAAGAATGGTGGCAGCAGGATCC 1431
Db 325 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysIysAsnGlyAlaGlnAlaGlySer 344
QY 1432 AATAACTCTGATCAGTTTCTTTCATGAGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491
Db 345 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 364
QY 1492 GCTGCTGGGAAGGGGAGCTATTTATGCCAAAAGCTCTCGTTGCTAACTGTGCCCT 1551
Db 365 AlaAlaGlyIysGlyGlyAlaIleTyrAlaIysIysLeuSerValAlaAsnCysGlyPro 384
QY 1552 GTACAAATTTTAAGGAATATCGCTAATGATGGTGGAGGATTTATTTAGGAGAACTGGA 1611
Db 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly 404
QY 1612 GAGCTCAGTTTATCTGCTGATTCGAGATATATTTTCGATGGGAATCTTAAAGAACAA 1671
Db 405 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuIysArgThr 424
QY 1672 GCCAAGAGAAATCGCCGATGTTAATGGCGTAACGTGTCTCTCAACAGCCATTTTCGATG 1731
Db 425 AlaIysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444
QY 1732 GCATCGGAGGAGGAAATAACGACATTAAGCAGCTAAAGCAGGCGCATCAGATTTCTTTAAT 1791
Db 445 GlySerGlyGlyIysIleThrThrLeuArgAlaIysAlaGlyHisGlnIleLeuPheAsn 464
QY 1792 GATCCCATCGAGATGGCAACCGAAATAACCGACCGAGCCAGCTTTTCCAAACTTCTTAAA 1851

465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484
1852 ATTAACCATGTAAGGATACACAGGCGATATTCTTTTGTCTAATGAAGCAGTACTTTG 1911
485 IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504
1912 TACCAAAATGTTACGATAGACGAAGGAGTCTTCTCGTGAAGGCAAAATATCA 1971
505 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 524
1972 GTGAATCTCTAAGTCACAGAGTGGAGCTGTGTATATGAGCTGGAGTCACATGGAT 2031
525 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 544
2032 TTCTGTAATCCACAAACACACACAGCTCTCGCTAATCACTAGTTGATCACGCTTTCC 2091
545 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 564
2092 AATCTGCATTTGCTCTTTCTTTCTTTGTTAGCAACAATGACGATTACGAATCTCTACC 2151
565 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 584
2152 AATCTCCAGCGCAAGATCTCATCTCCAGTCATTGTTAGCACAACTGCTGTTCTGTT 2211
585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 604
2212 ACAATTAGTGGCCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGGTATGAT 2271
605 ThrIleSerGlyProIlePhePheGluAspLeuAspThrAlaTyrAspArgTyrAsp 624
2272 TGGCTAGGTTCTAATCAAAAATCAATGCTCTGAAATTTACAGTTAGGACCTAAGCCCA 2331
625 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 644
2332 GCTAATGCCCATCGATTGACTCTAGGGAATGAGATGCCCTAAGTATGGCTATCAAGGA 2391
645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 664
2392 AGCTGGAAGCTTGGTGGGATCTTAATACAGCAAAATATGTCCTTACTCTGAAAGCT 2451
665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684
2452 ACATGGACTAAACCTGGGTATAAATCTCGGCTGAGCAGTACTTCTTTGTTCCAAAT 2511
685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704
2512 AGTTATCGGATTCATTTAGATATACATCTGCGCATTCAGCAATTCAGCAAGTGTG 2571
705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
2572 GATGGCGCTCTTATTTGTCGAGATTATGGGTTCTCGAGTTTCGAAATTTCTTCTATCAT 2631
725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPheTyrHis 744
2632 GACCGCGATGCTTTAGTCAGGATATCGGTATATTAGTGGGGTTATTCCTTAGAGACA 2691
745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
2692 AACTCTACTTTGGATCATCGATGTTGTCAGCATTTACCGAAGTATTGGTAGATCT 2751
765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
2752 AAGGATTATGATGTGTCCTCAATCATCATGCTTGCATAGGATCCGTTTATCTATCT 2811
785 LysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyrLeuSer 804
2812 ACCCAACAGCTTTATGTGGATCCTATTGTTGGAGATGCGTTTATCGCTAGCTAC 2871
805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824
2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCGAGGAGCGGATGTTGCT 2931
825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844

2932 TGGGATAATACTGTCTGCTGGCTGGAGAGATTGGAGCGGATACCGATTGTGATTACTCCA 2991
845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864
2992 TCTAAGCTCTATTGATGAGTTGCGTCTTTCTGTCGAAGCTGAGTTTCTTATGCGCAT 3051
865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884
3052 CATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCAATTCAGAGCGGACATCTCCTA 3111
885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
3112 AATCATCATGTTCTGTTGGAGTGAAGTTTGTATGATGTTCTTAGTACACATCTCTAATAA 3171
905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924
3172 TATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231
925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
3232 ACGTCTCTATCCCATCAAGAGACATGAGCAACAGATGCTCTTCTTCAATTTAGCAAGCATGGA 3291
945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964
3292 GTTCTGTTAGAGATCTATGATGTTCTTCTTAAACAAGTATATAGAAAGTATATGCGCAT 3351
965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984
3352 GGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTGAGTCAGGAGGATAGAGTC 3411
985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerIleVal 1004
3412 CGGTTCT 3417
1005 ArgPhe 1006
RESULT 24
US-10-872-155-190
; Sequence 190, Application US/10872155
; Publication No. US20040234536A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C10
; CURRENT APPLICATION NUMBER: US/10/872,155
; CURRENT FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 09/620,412
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 09/598,419
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/556,877
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/454,684
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 09/426,571
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 09/410,568
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/288,594
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 09/208,277
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-10-872-155-190


```
QY 2512 AGTTATCGGATCCATTTAGATATACATCTCGCATTCAGCAATTCAGCAAGTGTG 2571
Db 705 SerLeuTyrGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
QY 2572 GATGGCGCTCTATTGTCGAGGATATAGGGTTCTCGAGTTTCGAAATTTCTTCATCAT 2631
Db 725 AspGlyArgSerTyrCysArgGlyLeuTyrValSerGlyValSerAsnPhePheTyrHis 744
QY 2632 GACCGCATGCTTTAGTTCAGGATATCGGTATATTAGTGGGGTTATTCTTCAGGACA 2691
Db 745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
QY 2692 AACTCCTACTTTGGATCATCGATTTGGTCTAGCATTTACCAAGTATTTGGTAGATCT 2751
Db 765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
QY 2752 AAAGATTATGATGTGTGCTTCCATCATCATGCTTCATAGGATCGTTTATCTATCT 2811
Db 785 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 804
QY 2812 ACCCAACAGCTTTATGTGGATCCTATTGTTGCGGAGATGCGTTTATCCGTCTAGCTAC 2871
Db 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824
QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGGAGGCGATGTTGCT 2931
Db 825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844
QY 2932 TGGGATAATAACTGTCGCTGGAGATTCGAGCGGATTAACCATGTGATTAATCTCCA 2991
Db 845 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 864
QY 2992 TCTAAGCTCTATTGAATGAGTGTGCTTCTTCGTCGCAAGCTGAGTTTCTTATCCGAT 3051
Db 865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884
QY 3052 CATGAATCTTTTACAGAGAGGCGATCAAGCTCGGGCATTCAGAGCGGAGCATCTCCTA 3111
Db 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
QY 3112 AATCTATCAGTTCCTGTTGGAGTGAAGTTTGATCGATGCTTCAGTACACATCCTAATAA 3171
Db 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924
QY 3172 TATAGCTTTATCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTACTGAGACA 3231
Db 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
QY 3232 ACGCTCCTATCCCATCAGAGACATGGACAACAGATGCTTCATTTAGCAAGACATGGA 3291
Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964
QY 3292 GTTGTGGTTAGAGGATCTATGATGCTTCTCTAAAGTAATATAGAGTATATGCGCAT 3351
Db 965 ValValValAlaArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984
QY 3352 GAAAGATGATGATTCGAGATGCTTCTCGAGGCTATGTTTGGTGCAGGAAGTAGAGTC 3411
Db 985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004
QY 3412 CGGTTC 3417
Db 1005 ArgPhe 1006
```

RESULT 25

US-09-841-132-176

; Sequence 176, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841.132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: PASTSEQ for Windows Version 3.0/4.0

; SEQ ID NO 176

; LENGTH: 982

; TYPE: PRT

; ORGANISM: Chlamydia

; FEATURES:

; NAME/KEY: VARIANT

; LOCATION: (1)...(982)

; OTHER INFORMATION: xaa = Any Amino Acid

US-09-841-132-176

```
Alignment Scores:
Pred. No.: 0          Length: 982
Score: 5084.00        Matches: 978
Percent Similarity: 99.8%      Conservative: 2
Best Local Similarity: 99.6%    Mismatches: 2
Query Match: 64.5%             Indels: 0
DB: 3                         Gaps: 0
```

US-10-701-844-1 (1-4435) x US-09-841-132-176 (1-982)

```
QY 472 ATGGTTCCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATATCTGT 531
Db 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20
QY 532 ATAGAGATCCGAGTGGGACTACTGTTTTTCTCGCAGGAGATTAACTTAAAAATCTT 591
Db 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40
QY 592 GACAACTCTATTTCAGACTTTCCTTAAAGTTGTTTTGGAACTTATTAGGAGATTTTACT 651
Db 41 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60
QY 652 GTTTTAGGAGAGGACACTCGTTGACTTTCGAGAACATACGACTTCTACAAATGGGACA 711
Db 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTAGGGTTTAAAGAATTATCC 771
Db 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 100
QY 772 TTTTCCAATTGCAATTTCATTACTTCCCTGCTACTGCTGCTGCAACGACTTAATAAGGGTAGC 831
Db 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaIleThrAsnLysGlySer 120
QY 832 CAGACTCCGACGACAAACATCTACACCGCTAATGGTACTATTATTCTTAAACACATCTT 891
Db 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 140
QY 892 TTGTTACTCAATAATAGAGAAGTTCTCATCTATAGTAATTTAGTCTCTCGAGATGGGGA 951
Db 141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160
QY 952 GCTATAGATGCTTAAGAGCTTAACGGTTCAAGGAATTTAGCAAGCTTTGTGTCTTCCAGAA 1011
Db 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180
QY 1012 AATACTGCTCAAGCTGATGGGAGCTGTGTCAAGTAGTCACCAAGTTTCTCTGTATGGCT 1071
Db 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200
QY 1072 AACGAGGCTCCTATTGCTTTGTAGCGGAATGTTGAGGAGTAAAGGGGAGGAGTGTCT 1131
Db 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 220
QY 1132 GCTGTTTCAGGATGGGACAGGAGTGTCAATCATCTACTTCAACAGAGATCCAGTAGTA 1191
```


Db 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal 240
 QY 1192 AGTTTTTCCAGAAATCTCCGGTAGAGTTTGATGGAACTGAGCCGAGTAGAGAGGG 1251
 Db 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260
 QY 1252 ATTACTCTACGGAGAGCTGCTTCCGTAATATGGAATAACCTTGTCTCAACAAT 1311
 Db 261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsn 280
 QY 1312 GTTGCTTCTCTGTTTACATTTGCTGTAAGCAACCAAGTGGACAGGCTTCTAATACG 1371
 Db 281 ValAlaSerProValTyrIleAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300
 QY 1372 AGTAATAATTACGGAGATGGAGAGCTATCTCTGTAAGATGGTCCGACAGAGATCC 1431
 Db 301 SerAsnAsnTyrGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320
 QY 1432 AATACTCTGGATCAGTTCTCTTGTGATGGAGGGAGTAGTTTCTTTAGTAGCAATGTA 1491
 Db 321 AsnAsnSerGlySerValSerPheAspGlyGlyLysValPhePheSerSerAsnVal 340
 QY 1492 GCTGCTGGGAAAGGGGAGCTATTTATGCCAAAAGCTCTCGGTGCTAACTGTGCCCT 1551
 Db 341 AlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360
 QY 1552 GTA CAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGAATTTATTAGAGAACTTGG 1611
 Db 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 380
 QY 1612 GAGCTCAGTTTATCTGCTGATTTAGGATATTTATTTCCATCGGATCTTAAAGAAC 1671
 Db 381 GluLeuSerLeuSerAlaAspTyrGlyAspIlePheAspGlyAsnLeuLysArgThr 400
 QY 1672 GCCAAAGAGAAATGCTGCCGATGTTAATGGCGTAATCTGTCTCCTCAAGCCATTTCCGATG 1731
 Db 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420
 QY 1732 GGATCGGGAGGGAATAACGACATTAAGAGCTAAAGACGAGCGGCATCAGATCTCTTTAAT 1791
 Db 421 GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440
 QY 1792 GATCCCATCAGATGGCAACGGAATAACGACGAGCGAGCTCTCCAACTCTCAAAA 1851
 Db 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys 460
 QY 1852 ATTAACCATGCTGAAGGATACACAGGGATATCTGTTTCTAATGGAAGCAGTACTTTG 1911
 Db 461 IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
 QY 1912 TACCAAAATGTTACGATAGACAGCAAGGAGGATTTCTCTCGTGAAGGCAAAATTAACA 1971
 Db 481 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500
 QY 1972 GTGAATTTCTTAAGTACAGAGTGGAGCTGTATATGAGAGCTGGAGTACATGGAT 2031
 Db 501 ValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 520
 QY 2032 TTTGTAATCCACACACCAACAGCTCTCCGCTAATCAGTTGATCAGCTTCC 2091
 Db 521 PheValThrProGlnProProGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 540
 QY 2092 AATCTGCAATTTGCTCTCTTCTTCTTTGTTAGCAACCAATGAGTTACGAATCTCTTACC 2151
 Db 541 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProThr 560
 QY 2152 AATCTCAGCGCAGATTTCTATCTCTGATGAGTGGTACACACATGCTGGTCTGTT 2211
 Db 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580
 QY 2212 ACAATTAGTGGCTCTCTTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGAT 2271
 Db 581 ThrIleSerGlyProIlePhePheGluAspLeuAspThrAlaTyrAspArgTyrAsp 600

QY 2272 TGGCTAGGTTCTTAATCAAAAAATCAATGCTCTGAAATTTACAGTTAGGAGTAAGCCCCA 2331
 Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
 QY 2332 GCTAATGCCCCATCAGATTTGACTCTAGGAATAGATGCCCTAAGTATGGCTATCAGCA 2391
 Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 640
 QY 2392 ACCTGGAAGCTTGGCTGGGATCCTAATACAGCAAAATATGGTCTTATATCTCTGAAAGCT 2451
 Db 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660
 QY 2452 ACATCGACTAAAACCTGGGTATATCTGGGCTCGAGCGAGTAGCTTCTTTGGTTCAAAAT 2511
 Db 661 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
 QY 2512 AGTTATGGGGATCCATTTTAGATATATCGATCTGGCATTTCAGCAATTCACGCAAGTGG 2571
 Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
 QY 2572 GATGGGCTCTTATTTGTCGAGGATTTATGGGTTCTGAGATTTCGAATTTCTTCTATCAT 2631
 Db 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720
 QY 2632 GACCGCATGCTTTTAGTTCAGGATATCGGTATATTTAGTGGGGTTTATTCCTTTAGAGCA 2691
 Db 721 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740
 QY 2692 AACTCTCTCTTGGATCATCGATTTGTCTAGCATTTTCCGAAATTTTGTGTAGTACT 2751
 Db 741 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
 QY 2752 AAGATTATGATGTCGTTCCATCATCATCTGTCGATAGTCCGTTTATCTATCTATCT 2811
 Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
 QY 2812 ACCCAACAGCTTTATGTCGATCTTATTTGTCGAGATCGGTTTATCCGTGCTAGCTAC 2871
 Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800
 QY 2872 GGGTTTGGGAATCAGCATATGAAAACCTCATATACATTTTCAGAGAGCGATGTCGT 2931
 Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820
 QY 2932 TGGGATTAATAACTGCTGCTGAGAGATGGAGCGGATTCGATGTTGATTACTCCA 2991
 Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
 QY 2992 TCTAAGCTCTATTGATGATGTCGCTCTTCTGTCGCAAGCTCAGTTTCTTATGCCGAT 3051
 Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
 QY 3052 CATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTA 3111
 Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
 QY 3112 AATCTATCAGTTCTGTTGAGTCAAGTTTGATCGATGTTCTAGTACACATCCTAATAAA 3171
 Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
 QY 3172 TATAGCTTTATGGCGCTTATATCTGATGTCGCTTATTCGACCATCTCTGGTACTGAGACA 3231
 Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920
 QY 3232 ACGCTCTATCCCATCAAGAGACATGACCAACAGATCCCTTTCATTTAGCAAGACATGGA 3291
 Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940
 QY 3292 GTTGCTGTTAGAGGATCTATGATGCTTCTTAACAAGTATATAGAGTATATATGCGCAT 3351
 Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960

1792 GATCCATCGAGATGGCAACGAAATACACAGCCAGCGCAGCTCTCCAAACTTCTAAA 1851
Db AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 460
1852 ATTAAACGATGGTGAAGATACACAGGGGATATGTTTTGCTAAATGGAGCAGTACTTTG 1911
Db IleAsnAspGlyGluGlyThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
1912 TACCAAAATGTTACGATAGACAGAGAGATGTTCTTCGTGAAAGCAAAATATCA 1971
Db TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500
1972 GTCAATCTCTAAGTCACAGAGTGGAGCTCTGTATATGGAAGCTGGAGTACATGGAT 2031
Db ValAsnSerLeuSerGlnThrGlyGlySerLeuThrMetGluAlaGlySerThrLeuAsp 520
2032 TTTGTAACCTCCAAACCAACACACAGCCTCTCGCGCTAATCAGTTGATCAGCTTTCC 2091
Db PheValThrProGlnProGlnGlnProAlaAlaAsnGlnLeuIleThrLeuSer 540
2092 AATCTGCATTTGCTCTTCTTTGTTAGTAAACAAATGCGAGTTAGGAATCTCTACC 2151
Db AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560
2152 AATCTCCAGCGCAAGATCTCATCTCGCAGTCATTTGTTAGCACACTGCTGTTCTGTT 2211
Db AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580
2212 ACAATTAGTGGGCTATCTTTTGTAGGATTTTGGATGATACAGCTTATGATAGTATGAT 2271
Db ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaThrAspArgTyrAsp 600
2272 TGCTAGTGTCTAATCAAAATCAATGCTCCTGAAATACAGTTAGGACTAAGCCCCCA 2331
Db TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
2332 GCTAATGCCCATCAGATTGACTCTAGGAAATGAGTGCCTAAGTATGCTATCAGGA 2391
Db AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly 640
2392 AGCTGGAAGCTTCGGTGGATCTTAATACAGCAAAATAGTCTTATATCTCGAAAGCT 2451
Db SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660
2452 ACATGGACTAAACTGGGTATAATCTCGGCCTGAGCGAGTAGCTCTTTGGTTCCAAAT 2511
Db ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
2512 AGTTATGGGGATCCATTTAGATATACGATCTCGGCATTCAGCAATTCAGCAAGTGTG 2571
Db SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
2572 GATGGGCGCTTATCTCGAGGATATGTTGTTCTCGAGTTTCGAAATTTCTTCTATCAT 2631
Db AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720
2632 GACCGCATCTTTAGTTCAGGGATATCGGTATATAGTGGGGTATTCCTTAGGAGCA 2691
Db AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740
2692 AACTCTACTTTGGATCATGATGTTGGTGTAGCATTTACGAGTATTTGGTAGTCT 2751
Db AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
2752 AAGATTATGTAGTGTGCTGTTCAATCATCATCTGTCATAGATCCGTTTATCTATCT 2811
Db LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
2812 ACCCAACAAGCTTTATGTGGATCTTATTTGTTTCGGAGATGCGTTTATCCGTGCTAGTAC 2871
Db ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800

2872 GGGTTTGGGAATCAGCATATCAAAACCTCATATACATTTGCGAGGAGAGCGATTTCTGT 2931
Db GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820
2932 TGGGATAATACTGCTGCTCGGAGATTCGAGCGGATTCGAGCGGATTCACGATTTGTATCTCA 2991
Db TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
2992 TCTAAGCTCTATTTGAATGAGTTGCTCTCTTTCTGTCGCAAGCTGAGTTTTCTTATCCGAT 3051
Db SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
3052 CATGAATCTTTTACAGAGAGGCGATCAAGCTCGGGCATTCAGAGCGGACATCTCTCA 3111
Db HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
3112 AATCTATCAGTCTCTGTTGGAGTGAAGTTGATCGATGTTCTAGTACACATCTTAATAA 3171
Db AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
3172 TATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTACTGAGACA 3231
Db TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920
3232 ACGCTCTATCCATCAAGAGACATGGACAACAGATGCGCTTTTCATTTAGCAAGACATGGA 3291
Db ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940
3292 GTTGTGTTAGGATCTATGATGCTCTCTTAACAAGTAATATAGACTATATGCGCAT 3351
Db ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960
3352 GGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGGTGCAGGAAGTAGATGC 3411
Db GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980
3412 CGGTTTC 3417
Db 981 ***Phe 982
RESULT 27
US-10-197-220-169
; Sequence 169, Application US/10197220
; Publication No. US20050084499A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCES: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 169
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Chlamydia
US-10-197-220-169
Alignment Scores: 8 92e-288 Length: 670
Pred. No.: 3336.50 Matches: 651
Percent Similarity: 98.8% Conservative: 11
Best Local Similarity: 97.2% Mismatches: 7
Query Match: 42.3% Indels: 1
DB: 5 Gaps: 1
US-10-701-844-1 (1-4435) x US-10-197-220-169 (1-670)
QY 463 GCAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGATCATTTCCC 522

Db 1 AlaGluIleMetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPhePro 20
Qy 523 TATACGTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTCTCGCAGGAGAGTTAAACATTA 582
Db 21 TyrThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu 40
Qy 583 AAAATCTTGACAACTTCTATTGCAGCTTGGCTTTAACTGTTGTTGGGAACTATTAGGG 642
Db 41 LysAsnLeuAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly 60
Qy 643 AGTTTTACTGTTTATAGGAGGAGACACTGCTTGACTTTTCGAAACATACGAGCTTCTACA 702
Db 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80
Qy 703 AATGGGCGAGCTCTAAGTAATAGCCGCTGCTGATGGACTGTTTACTATTAGAGGTTTAAA 762
Db 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100
Qy 763 GNAATTACTTTTCCAAATGCAATTCATTACTTGGCGTACTGCCCTGCTCAACGACTAAT 822
Db 101 GluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsn 120
Qy 823 AAGGGTAGCCAGACTCCGAGCAACATCTACACCGTCTAAATGGTACTATTATTCTAAA 882
Db 121 AsnGlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLys 140
Qy 883 ACAGATCTTTGTTTACTCAATATAGAAAGTTCTCATTTCTATPAGTAATTTAGTCTCGGA 942
Db 141 ThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGly 160
Qy 943 GATGGGGAGCTATAGATCTTAAGACTTAAGCTTCAAGGTTCAAGGAATTAGCAGCTTGTGTC 1002
Db 161 AspGlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysVal 180
Qy 1003 TTCCAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCAACAGTTTCTCT 1062
Db 181 PheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSer 200
Qy 1063 GCTATGGCTAACGAGGCTCTATTGCTTTGTAGCGAATGTTGCGAGGAGTAAGAGGGGA 1122
Db 201 AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly 220
Qy 1123 GGGATGCTGCTGTTCCAGATGGCGAGGAGGTGTCATCTACTTCAACAGAGAT 1182
Db 221 GlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAsp 240
Qy 1183 CCAGTAGTAAGTTTTCCAGAAATACTGCGGTAGAGTTTGATGGGAACGTAGCCGAGTA 1242
Db 241 ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal 260
Qy 1243 GGAGGAGGATTTACTCTACGGGAACGTTGCTTTCTCGAATAATGGAAAAACCTTGTGTT 1302
Db 261 GlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPhe 280
Qy 1303 CTCACAAATGTTGCTTCTCTGTTTACATTTGCTTACGACCAACCAAGTCGACAGGCT 1362
Db 281 LeuAsnAsnValAlaSerProValTyrIleAlaAlaGluGlnProThrAsnGlyGlnAla 300
Qy 1363 TCTAATACGAGTAATAATACGAGATGGAGAGCTATCTCTGTAAGAATGTCGCGCAA 1422
Db 301 SerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGln 320
Qy 1423 ---GCAGGATCCAATACTCTCGATCAGTTTCTTTGATGGAGAGGAGTACTGTTTCTTT 1479
Db 321 AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValPhePhe 340
Qy 1480 AGTAGCAATGTAAGTCTGGGAAAGGGGAGCTATTATTATGCCAAAAGCTCTCGTTGCT 1539
Db 341 SerSerAsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAla 360
Qy 1540 AACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAAATGATGGTGGAGCGATTTATTTA 1599
Db 361 AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyrLeu 380

Qy 1600 GGAGATCTGGAGAGCTCAGTTTATCTGCTGATATTCGAGATATATTTTCGATGGGAT 1659
Db 381 GlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsn 400
Qy 1660 CTTAAAAGAACAGCAAGAGAAATGTCGCCGATGTTAATGCGTAACGTGTCTCTCACAA 1719
Db 401 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 420
Qy 1720 GCCATTTTCGATGGGATCGGGAGGAAAATAACGACATTAAAGAGCTAAACGAGGCGATCAG 1779
Db 421 AlaIleSerMetGlySerGlyGlyIleThrThrLeuArgAlaLysAlaGlyHisGln 440
Qy 1780 ATTCTCTTTAATGATCCCATCGAGATGGCAAAACGGAATAACCCAGCCAGCGAGCTTCC 1839
Db 441 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 460
Qy 1840 AAACCTTTCAAAAATTAACGATGGTGAAGGATACACAGGGGATATGTTTTTGTCTAATGA 1899
Db 461 GluProLeuLysIleAsnAspGlyGlyTyrThrGlyAspIleValPheAlaAsnGly 480
Qy 1900 AGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGGAAGATTGTTCTTGTGTAAG 1959
Db 481 AsnSerThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLys 500
Qy 1960 GCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGCTGGAGTCTGTATATCGAAGCTGG 2019
Db 501 AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGly 520
Qy 2020 AGTACATCGGATTTTGTAACTCCACACACACAGCCTCTCTGCGCTAAATCAGTTG 2079
Db 521 SerThrLeuAspPheValThrProGlnProGlnGlnProAlaAlaAsnGlnLeu 540
Qy 2080 ATCAGCTTTTCCAATCTGCATTTGCTCTTTCTTTTGTAGCAAAACAATGCAGTTACG 2139
Db 541 IleThrLeuSerAsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThr 560
Qy 2140 AATCTCTCTCAATCTCCAGCGCAAGATTCTCATCTCTGCACTGTTGTTAGCAACT 2199
Db 561 AsnProThrAsnProAlaGlnAspSerHisProAlaIleIleGlySerThrThr 580
Qy 2200 GCTGTTCTGTTTACAAATTTAGTGGGCTATCTTTTGTAGGATTTGATGATACAGCTTAT 2259
Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAspLeuAspThrAlaTyr 600
Qy 2260 GATAGTATGATTTGCTAGTTTCTAATCAAAAATCAATGTCTCTCAAAATACAGTTAGGG 2319
Db 601 AspArgTyrAspTyrLeuGlySerAsnGlnLysIleAspValLeuLysLeuGlnLeuGly 620
Qy 2320 ACTAAGCCCCAGCTAATGCCCCCATCAGATTGACTCTAGGGAATGAGATGCTAAGTAT 2379
Db 621 ThrGlnProSerAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyr 640
Qy 2380 GGTATCAAGGAGCTGGAAGCTTCGGTGGGATCTCTAATACAGCAAAATAATGCTCTTAT 2439
Db 641 GlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsnThrAlaAsnAsnGlyProTyr 660
Qy 2440 ACTCGAAGCTACATGGAATACTGGG 2469
Db 661 ThrLeuLysAlaThrTyrThrLysThrGly 670

RESULT 28

US-11-109-468-169
; Sequence 169, Application US/11109468
; Publication No. US20050232941A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C4

```

; CURRENT APPLICATION NUMBER: US/11/109,468
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US 10/197,220
; FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 10/007,693
; FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 10/012,256
; FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 09/841,260
; FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/219,752
; FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/198,853
; FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 169
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Chlamydia
US-11-109-468-169

```

```

Alignment Scores:
Pred. No.: 8,92e-288 Length: 670
Score: 3336.50 Matches: 651
Percent Similarity: 98.8% Conservative: 11
Best Local Similarity: 97.2% Mismatches: 7
Query Match: 42.3% Indels: 1
DB: Gaps: 1

```

US-10-701-844-1 (1-4435) x US-11-109-468-169 (1-670)

```

QY 463 GCAGAAATCATGTTCTCTCAGGAATTTAGGATGGGAGACGTTAATCTGATCAATTTCCC 522
DB 1 AlaGluMetIleProGlnGlyIleTyAspGlyGluThrLeuThrValSerPhePro 20
QY 523 TATACGTGTATAGGACATCCGAGTGGGACTACTGTTTCTCTCAGGAGAGTTAAACATTA 582
DB 21 TyThrValIleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeu 40
QY 583 AAAAACTTTGCACAAATCTATTGACGCTTTGCGCTTTTAAAGTTGTTTGGGAACTATTAGGG 642
DB 41 LysAsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGly 60
QY 643 AGTTTACTGTTTGGGAGAGACACTGTTGACCTTTTCGAGACATACGACTCTACA 702
DB 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80
QY 703 AATGGGCGACTTAAGTAAATAGCTGCTGATGAGTGTATTACTATTGAGGGTTTAAA 762
DB 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100
QY 763 GAATTATCTTTTCCAAATTCGAATTCATTAATCTGCGTACTGCTGCTCAACGACTAAT 822
DB 101 GluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaIleThrAsn 120
QY 823 AAGGTAGCCAGACTCCGACGACAACTATCAACCGTCTAATGGTACTATTATTCTAAA 882
DB 121 AsnGlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLys 140
QY 883 ACAGATCTTTGTACTCAATATGAGAGTTCTCTCAATCTATAGTAATTTAGTCTCTGGA 942
DB 141 ThrAspLeuLeuLeuLeuAsnGlyPheSerPheTySerAsnLeuValSerGly 160
QY 943 GATGGGCGAGCTATAGATGCTTAAGAGCTTAACGGTTTCAGGAATTAACGAGCTTTGTGTC 1002
DB 161 AspGlyAlaIleAlaLeuSerLeuThrValGlnGlyIleSerLysLeuCysVal 180
QY 1003 TTCCAAGAAATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCAACAGTTTCTCT 1062
DB 181 PheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSer 200
QY 1063 GCTATGCTTAACGAGGCTCTATTGCTTTGTAGCGAATGTTGCGAGGTAAGAGGGGGA 1122

```

```

DB 201 AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly 220
QY 1123 GGGATTGCTGCTGTTTCCAGGATGGGAGGAGGTGTCATCTATCTTACACAGAAGAT 1182
DB 221 GlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerThrSerThrGluAsp 240
QY 1183 CCAGTAGTAAGTTTTCAGAAATCTACCGGTAGAGTTGATGGGAACTAGCCCGAGTA 1242
DB 241 ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal 260
QY 1243 GGAGGAGGATTTACTCTACGGGAACTTCTCTTCTCTTCTCTCTCTCTCTCTCTCTCT 1302
DB 261 GlyGlyGlyIleTySerTyGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPhe 280
QY 1303 CTCACAAATGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1362
DB 281 LeuAsnAsnValAlaSerProValTyIleAlaGluGlnProThrAsnGlyGlnAla 300
QY 1363 TCTAATACGAGTAATTAATACGAGATGGAGAGGTATCTCTCTCTCTCTCTCTCTCT 1422
DB 301 SerAsnThrSerAspAsnTyGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGln 320
QY 1423 ---GCAGATCCAAATACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1479
DB 321 AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePhe 340
QY 1480 AGTAGCAATGATGCTGCTGGGAAAGGGGAGCTATTATTCGCAAAAAGCTCTCGGTGCT 1539
DB 341 SerSerAsnValAlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAla 360
QY 1540 AACTGTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGATGGTGGACGATTTATTTA 1599
DB 361 AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyLeu 380
QY 1600 GGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTAATGAGAGATATTTTTCATGCGAAT 1659
DB 381 GlyLysSerGlyGluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsn 400
QY 1660 CTTTAAAGAACACCAAGAGAAATGCTGCGGATGTTAAATGGCGTAACTGTGCTCTCACAA 1719
DB 401 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 420
QY 1720 GCATTTTCATGGGATCGGGAGGGAATAATACACATTAAGAGCTAAAGCAGGCGATCAG 1779
DB 421 AlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln 440
QY 1780 ATTCTCTTAAATGATCCCATCGAGATGGCAACCGGAATAACACGACGCGGAGTCTTCC 1839
DB 441 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 460
QY 1840 AAACCTCTTAAATTAACGATGCTGAAGGATACACAGGGGATATGCTTTTGTCTAATGGA 1899
DB 461 GluProLeuLysIleAsnAspGlyGlyGlyTyThrGlyAspIleValPheAlaAsnGly 480
QY 1900 AGCAGTACTTTGTACCAAAATGTTACGATAGACAGGAGGAGGATGTTCTTCTGTGAAAAG 1959
DB 481 AsnSerThrLeuTyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLys 500
QY 1960 GCAAAAATTATCAGTGAATTTCTTAAGTCACAGAGTGGGAGTGTGTATATGGAAGCTGGG 2019
DB 501 AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyMetGluAlaGly 520
QY 2020 AGTACATGGATTTTGTAACTCCACACCAACCAACGCTCTCTCTCTCTCTCTCTCTCT 2079
DB 521 SerThrLeuAspPheValThrProGlnProProGlnGlnProProAlaAlaAsnGlnLeu 540
QY 2080 ATCACGCTTTCCAATCTGCAATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2139
DB 541 IleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThr 560
QY 2140 AATCTCTCTACCAATCTCTCCAGCGCAAGATTCTCATCTCTCAGTCAATTTGTAGACA 2199
DB 561 AsnProProThrAsnProProAlaGlnAspSerHisProAlaIleIleGlySerThrThr 580

```

QY 2200 CTGCTGTTCTGTACAAATTAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTAT 2259
Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAsePLeuAsePThrAlaTy 600
QY 2260 GATAGGTATGATGGCTAGGTTCTTAATCAAAAATCAATGTCCTGAAATACAGTTAGGG 2319
Db 601 AspArgTy 620
QY 2320 ACTAAGCCCCAGCTAAATGCCCATCAGATTGACTCTAGGGAATGAGATGCTAAAGTAT 2379
Db 621 ThrGlnProSerAlaAseAlaProSerAspLeuThrLeuGlyAseGluWetProLysTy 640
QY 2380 GGCTATCAAGGAAGCTGGAAGCTCGTGGGATCTTAATACAGCAAAATAATGCTCTTAT 2439
Db 641 GlyTy 660
QY 2440 ACTCTGAAGCTACATGGACTAAACTGGG 2469
Db 661 ThrLeuTy 670

RESULT 29

US-10-701-844-17
; Sequence 17, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-399
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-17

Alignment Scores:

Pred. No.: 6,71e-218 Length: 505
Score: 2552.00 Matches: 505
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.4% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-701-844-17 (1-505)

QY 466 GAAATCATGTTCTCTCAAGGAATATTCAGATGGGAGACCTTAACCTGTATCATTTCCCTAT 525
Db 1 GluIleMetValProGlnGlyIleTy 20
QY 526 ACTGTTATAGGATCCGAGTGGGACTACTGTTTCTCGCAGGAGAGTTAATCAATAAA 585
Db 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAAATCTTATTCAGCTTTCCTTTAAGTTGTTTGGGAACTTATTAGGGAGT 645
Db 41 AsnLeuAseAseSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60
QY 646 TTTACTGTTTATGGGAGAGACATCTGTTGACTTCCTCGAAGACATACGAGCTTCTCAAT 705
Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAseIleArgThrSerThrAsn 80
QY 706 GGGGCGAGCTTAAGTAATAGCGCTCTGATGACTGTTTACTATTGAGGGTTTAAAGAA 765
Db 81 GlyAlaAlaLeuSerAseSerAlaAlaAsePLeuPheThrIleGlyPheLysGlu 100

QY 766 TTATCTTTTCCAAATTGCAATTCAATTACTTGCCCTACTGCTGCTGCAACGACTAATAAG 825
Db 101 LeuSerPheSerAseCysAseSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
QY 826 GGTAGCCAGACTCCGACGACAACTACACCGCTCAATGGTACTATTATTCTTAAACA 885
Db 121 GlySerGlnThrProThrThrSerThrProSerAseGlyThrIleTySerLysThr 140
QY 886 GATCTTTTGTACTCAATAATGAGAAGTTCTATTCTATAGTAATTTAGTCTCTGGAGAT 945
Db 141 AspLeuLeuLeuAseAseGluLysPheSerPheTySerAseLeuValSerGlyAse 160
QY 946 GGGGAGCTATAGATCTAAGAGCTTAACGTTTCAAGGAATTAGCAAGCTTTGTCTTC 1005
Db 161 GlyGlyAlaIleAsePAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
QY 1006 CAAGAAATACCTGCTCAAGCTGATGGGAGCTGTCAAGTAGTCACCAGTTTCTCTGCT 1065
Db 181 GlnGluAseThrAlaGlnAlaAsePglyAlaCysGlnValValThrSerPheSerAla 200
QY 1066 ATGCTTAACGAGGCTCTTATTGCTTGTAGCGAATGTTGAGGAGTAAAGGGGAGGG 1125
Db 201 MetAlaAseGluAlaProIleAlaPheValAlaAseValAlaGlyValArgGlyGly 220
QY 1126 ATTGCTGCTTTCAGGATGGCAGAGGAGTGTCATCATCTTCAACAGAGATCCA 1185
Db 221 IleAlaAlaValGlnAsePAlaGlnGlyValSerSerSerThrThrGluAsePro 240
QY 1186 GTAGTAAAGTTTCCAGAAATACCTGCGTACAGTTTGTATGGGAGCTAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerAseThrAlaValGluPheAsePAlaAseValAlaAseValGly 260
QY 1246 GGAGGATTTACTCTACGGCAAGCTGCTTCTCTGAATAATGGAATAACCTTGTCTC 1305
Db 261 GlyGlyIleTySerTyGlyAseValAlaPheLeuAseAseGlyLysThrPheLeu 280
QY 1306 AACAAATGTTCTCTCTGTTTACATGCTGCTAGCAACCAACAAAGTGGAGAGCTTCT 1365
Db 281 AseAseValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAse 300
QY 1366 AATACGAGTAATAATACGAGATGGAGGAGCTCTTCTGTAAGAAATGGTGGCAAGCA 1425
Db 301 AseThrSerAseAseTyGlyAsePglyAlaIlePheCysLysAseGlyAlaGlnAla 320
QY 1426 GGATCCATAACTCTGATCAAGTTTCTTGTATGGAGGAGTAAAGTTTCTTTAGTAGC 1485
Db 321 GlySerAseAseSerGlySerValSerPheAsePglyGlyGlyValValPhePheSer 340
QY 1486 AATGATGCTGCTGGGAAAGGGGAGCTATTTATGCAAAAAGCTCTCGGTTGCTAACTGT 1545
Db 341 AseValAlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAlaAseCys 360
QY 1546 GGCCCTGTACAAATTTTAAAGGAATATCCGTAATGATGGTGGAGCGATTTATTAGAGAA 1605
Db 361 GlyProValGlnPheLeuAseAseIleAlaAseAsePglyAlaIleTyLeuGlyGlu 380
QY 1606 TCTGAGAGCTCAGTTTATCTGCTATGAGATATTTTTCGATGGGAATCTTAAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAseTyGlyAsePglyIlePheAsePglyAseLeuLys 400
QY 1666 AGAACAGCAAGAGAGATGCTGCCGATGTTAATGGCGTAACTGTCTCTCACAGCCATT 1725
Db 401 ArgThrAlaLysGluAseAlaAseValAseGlyValThrValSerSerGlnAlaIle 420
QY 1726 TCGATGGGATCGGAGGAGAAATACGACTTAAGAGCTAAAGAGGAGCATCAGATCTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
QY 1786 TTTAATGATCCCATCGAGATGGCAACGGAATAACAGCCAGCCAGCTCTTCCAACTT 1845
Db 441 PheAseAseProIleGluMetAlaAseGlyAseAseGlnProAlaGlnSerSerLysLeu 460


```

QY 1846 CTAAATTAACGATGCTGAAGTACACAGGGGATATGTTTGTCTAATGGAAGCAGT 1905
DB 461 LeuLysIleAsnAspGlyGluGlyThrGlyAspIleValPheAlaAsnGlySerSer 480
QY 1906 ACTTTGTACCAAAATGTTACGATAGACGAGCAAGGAGATTGTTCTTGTGAAAAAGCAAAA 1965
DB 481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
QY 1966 TTATCATGTAATCTT 1980
DB 501 LeuSerValAsnSer 505

RESULT 30
US-10-766-711-17
; Sequence 17, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-17

Alignment Scores:
Pred. No.: 6,71e-218 Length: 505
Score: 2552.00 Matches: 505
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.4% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-766-711-17 (1-505)
QY 466 GAAATCATGTTCTCTCAGGAATTTAGATGGGAGACGTTAACTGATCATTTCCCTAT 525
DB 1 GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
QY 526 ACTGTTATAGGATCCGAGTGGGACTACTGTTTCTTCAGGAGAGTTAACTAAATAA 585
DB 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAAATCTATTGACGCTTTGCCCTTTTAAAGTTGTTTGGGAACTTTAGGGAGT 645
DB 41 AsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuLys 60
QY 646 TTTACTGTTTGGGAGAGACACTCGTTGACTTTCGAGACATACCGACTTCTACAAAT 705
DB 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
QY 706 GGGGACGCTTAAGTAAATAGCGCTGATGAGCTGTTTACTATTGAGGGTTTAAAGAA 765
DB 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100
QY 766 TTATCTTTTCCAAATTCGAATTCATTACTTTGCCGCTACTGCTGCTGCAACGACTAATAAG 825
DB 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
QY 826 GGTAGCCAGACTCCGACGACAACTCTACACCGTCTAATGCTACTATTATTCTTAAACA 885
DB 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLysThr 140
QY 886 GATCTTTTGTACTCAATATAGAGACTTCTCATTTAGTAAATTTAGTCTCTGGAGAT 945

```

```

DB 141 AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160
QY 946 GGGGAGCTATAGATGCTAAGAGCTTAAAGGTTCAAGCAATAGCAAGCTTTGTCTCTTC 1005
DB 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
QY 1006 CAAGAAAATACTGCTCAAGCTGATGGGAGCTTCTCAAGTAGTACCAAGTTTCTCTGCT 1065
DB 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAla 200
QY 1066 ATGGCTAACAGGCTCTTATTCCTTCTAGCGAATGTTGACGAGTAGTAAAGGGGAGGG 1125
DB 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
QY 1126 ATTTGCTGCTTTCAGGATGGGAGGAGGAGTGTCTATCTACTTCTCAACAGAGATCCA 1185
DB 221 IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspPro 240
QY 1186 GTAGTAAGTTTTCAGAAATATCTCGGTAGATTGTTGAGGAACTAGCCCGAGTAGGA 1245
DB 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
QY 1246 GGAGGATTACTCTCTACGGGAACGTTGCTTCTCAATAATGGAATAACCTTGTCTCTC 1305
DB 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
QY 1306 AACAAATGTTGCTTCTCTGTTTACATTTGCTGCTAAAGCAACCAAGTGCACAGGCTTCT 1365
DB 281 AsnAsnValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
QY 1366 AATACGAGTAATAATTACGGAGATGGAGGAGTATCTTCTGTAAGAAATGGTGGCAAGCA 1425
DB 301 AsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320
QY 1426 GGATCCAATACTCTGGATCAGTTCTTCTTGTAGAGGAGGAGTACTTCTTTTACTAGC 1485
DB 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340
QY 1486 AATGTAGCTGCTGGGAAAGGGGAGCTATTATTCACAAAAGCTCTCGGTGCTACTCTGT 1545
DB 341 AsnValAlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCys 360
QY 1546 GGCCCTGTACAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGAATTTATTAGGAGAA 1605
DB 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380
QY 1606 TCTGGAGAGCTCAGTTTACTGCTGATTTATGGAGATATTATTTTCATGGGAATCTTAAA 1665
DB 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400
QY 1666 AGACACAGCAAGAGAAATGCTCCGATGTTAATGGCGTAACTGCTGCTCCTCACAAGCCATT 1725
DB 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420
QY 1726 TCCATCGGATCGGGAGGAAAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCTC 1785
DB 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
QY 1786 TTTAATGATCCCATCCGAGATGGCAACGGAATAACAGCAGGCGGAGCTCTTCCAACTT 1845
DB 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
QY 1846 CTAAATAATTAACTGTTGAGGATACACAGGGGATATGTTTGTGTAATGGAAGCAGT 1905
DB 461 LeuLysIleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480
QY 1906 ACTTTGTACCAAAATGTTACGATAGACGAGGAGGAGTGTCTTCTGCTGAAAAAGCAAAA 1965
DB 481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
QY 1966 TTATCATGTAATCTT 1980
DB 501 LeuSerValAsnSer 505

```



```
RESULT 31
US-10-931-779-17
; Sequence 17, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-17

Alignment Scores:
Pred. No.: 6 71e-218 Length: 505
Score: 2552.00 Matches: 505
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.4% Indels: 0
Dbs: 5 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-17 (1-505)
QY 466 GAAATCATGTTCTCAAGGAATTTACGATGGGAGAGCTTAATGTATCATTTCCCTAT 525
Db 1 GluileMetValProGlnGlyIleTyraSpGlyGluThrLeuThrValSerPheProTy 20
QY 526 ACTGTTATAGGAGATCCGAGTGGGACTACTGTTTCTCGAGGAGAGTTAACTTAAA 585
Db 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu 40
QY 586 AATCTTGCAATTTCTATTCAGCTTTGCTTAAAGTTGTTTGGGAACTTATTAGGAGT 645
Db 41 AsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60
QY 646 TTTACTGTTTATAGGAGAGACACTCGTTGACTTTTCGAGAACATACGACTTCTCAAA 705
Db 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
QY 706 GGGGCGAGCTTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGGTTTTAAGNA 765
Db 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100
QY 766 TTATCTCTTTTCCAATTGCAATTCATTACTTGCCTGCTGCTGCAACGACTAATAAG 825
Db 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
QY 826 GGTAGCCAGACTCCGACGACATCTACACCGTCTAATGGTACTATTATTCTAAACA 885
Db 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTyraSerLysThr 140
QY 886 GATCTTTTGTACTCTAATATAGAAAGTTCTCATCTATAGTAATTTAGTCTCTGAGAT 945
Db 141 AspLeuLeuLeuAsnAsnGlnLysPheSerPheTyraSerAsnLeuValSerGlyAsp 160
QY 946 GGGGAGCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTC 1005
Db 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
QY 1006 CAAGAAATPACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTACCAAGTTTCTGCT 1065
Db 181 GlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSerAla 200
QY 1066 ATGGCTAACGAGCTCTCTATTGCTTTGTAGCGAATGTTGCGAGGAGTAAGAGGGGAGG 1125
```

```
Db 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
QY 1126 ATTCTGCTGTTAGGATGGGACGAGGAGTGTATCATCTACTTCAACAGAGATCCA 1185
Db 221 IleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerThrThrGluAspPro 240
QY 1186 GTAGTAAGTTTTTCCAGAAATACCTCGGTAGAGTTGATGGACGTAGCCGAGTAGGA 1245
Db 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
QY 1246 GGAGGATTTTACTCTACGGGAACGTTGCTTCTGAATAATGAAAAACCTTGTTCCTC 1305
Db 261 GlyGlyIleTyraSerTyraGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
QY 1306 AACAAATGTTGCTTCTCTGTTTACATTGCTCTAAGCAACCAACAAAGTGGAGAGCTTCT 1365
Db 281 AsnAsnValAlaSerProValTyraIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
QY 1366 AATACGAGTAATAATTACGGAGATGGAGGAGCTATCTCTTAAGATGGTGGCAAGCA 1425
Db 301 AsnThrSerAsnAsnTyraGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320
QY 1426 GGATCCAATAACTCTGGATCAGTTTCTTTCATCGAGAGGAGTAGTTTCTTTAGTAGC 1485
Db 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340
QY 1486 AATGTAGCTGCTGGGAAAGGGGAGCTATTATTATCCAAAAAGCTCTCGTTGCTAACTGT 1545
Db 341 AsnValAlaAlaGlyLysGlyAlaIleTyraLysLysLeuSerValAlaAsnCys 360
QY 1546 GGCCCTGTACAAATTTTAAAGGAATATCGCTAATGATGCTGGAGGAGTTATTATTAGGAA 1605
Db 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyraLeuGlyGlu 380
QY 1606 TCTGAGAGCTCAGTTTATCTGCTGATTATCGAGATATATTTTCGATGGGAATCTTAAA 1665
Db 381 SerGlyGluLeuSerLeuSerAlaAspTyraGlyAspIleIlePheAspGlyAsnLeuLys 400
QY 1666 AGAACAGCCAAAGAGAATGCTGCCGATTTAATGGCGTAATGCTGCTCTCAAGCCATT 1725
Db 401 ArgThrAlaLysGluAsnAlaAlaAspValAlaSerValThrValSerSerGlnAlaIle 420
QY 1726 TCGATGGATCGGAGGGGAAAAATAACGACATTAGAGCTAAAGCGGCGCATCAGATTCTC 1785
Db 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
QY 1786 TTTAATGATCCCATCGAGATGCAAAACGAAATAACAGCCAGCGCAGCTTTCCAAACCTT 1845
Db 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
QY 1846 CTAATAATTAACGATGGTGAAGGATACACAGGGGATATTGTTTTGCTTAATGGAAGCAGT 1905
Db 461 LeuLysIleAsnAspGlyGlyTyraThrGlyAspIleValPheAlaAsnGlySerSer 480
QY 1906 ACTTTGTACCAAAATGTTACGATAGCAAGAGGATGTTCTTCGTGAAAAGCCAAA 1965
Db 481 ThrLeuTyraGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
QY 1966 TTATCAGTGAATCTT 1980
Db 501 LeuSerValAsnSer 505
```

RESULT 32

```
US-10-701-844-36
; Sequence 36, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
```

```
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-36

Alignment Scores:
Pred. No.: 6,52e-200 Length: 458
Score: 2350.00 Matches: 458
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.8% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-701-844-36 (1-458)
QY 1030 GGGGAGCTGTCTAGTAGTACCAAGTTTCTCTGCTATGGCTAACGAGGCTCTATTGCC 1089
DB 1 GlyGlyAlaCyGlnValThrSerPheSerAlaMetAlaAsnGluAlaProileAla 20
QY 1090 TTTGTAGCGAATGTTGCGAGTAGTAGAGGGGAGGGATGCTGCTTCAGGATGGGAG 1149
DB 21 PheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGln 40
QY 1150 CAGGAGGTGCATCATCTACTTCCACAGAGATCCAGTAGTAAGTTTCCAGAAATACT 1209
DB 41 GlnGlyValSerSerThrSerThrGluAspProValValSerPheSerArgAsnThr 60
QY 1210 GCGGTAGAGTTTGTAGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAC 1269
DB 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleTySerTyArgAsn 80
QY 1270 GTTGCTTCTCGAATAAATGGAACCTGTTTCTCAACATGTTGCTTCTCTCTTTAC 1329
DB 81 ValAlaPheLeuAsnAsnGlyIleThrLeuPheLeuAsnAsnValAlaSerProValTy 100
QY 1330 ATTGCTGTACACCAACAGTAGGACAGGCTCTTAATACAGTAATAATACCGAGAT 1389
DB 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyArgAsp 120
QY 1390 GGAGGAGCTATCTCTGTAAGAATGGTGGCAAGCAGGATCCAATAACTCTGGATCAGTT 1449
DB 121 GlyGlyAlaIlePheCylAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140
QY 1450 TCTTTGTAGGAGGAGTAGTCTTTTCTTTAGTACAAATGTAGTCTGGGAAAGGGGA 1509
DB 141 SerPheAspGlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyIleGly 160
QY 1510 GCTATTTATCCAAAAGCTCTCGTTGCTGTAAGTGGCCCTGTAACATTTTAAAGGAT 1569
DB 161 AlaIleTyAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180
QY 1570 ATCGCTTAATGATGGTGGAGCGATTATTTTAGGAGATCTGGAGAGCTCAGTTTATCTGCT 1629
DB 181 IleAlaAsnAspGlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAla 200
QY 1630 GATTATGGAGATATTTATTTTCGATGGGAATCTTTAAAGAACAGCCAAAGAGATGCTGCC 1689
DB 201 AspTyArgAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla 220
QY 1690 GATGTTAATGGGTAACTGTGCTTCAAGCCATTTCCATGGATCGGAGGAGGAATA 1749
DB 221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIle 240
QY 1750 ACGACATTAAAGAGTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCA 1809
```

```
DB 241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProileGluMetAla 260
QY 1810 AACGAAATAACACGAGCCAGCGCAGTCTTCCAAACTTCTAAATAATTAACGATGGTGAAGGA 1869
DB 261 AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGly 280
QY 1870 TACACAGGGGATATGTTTGTCTTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATA 1929
DB 281 TyThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyArgAsnValThrIle 300
QY 1930 GAGCAAGGAGGATGTTCTCTGCGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAG 1989
DB 301 GluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGln 320
QY 1990 ACAGTGGGAGTCTGTATATGAAAGCTGGGAGTAGTACATGGGATTTTGTAACTCCCAACCA 2049
DB 321 ThrGlySerLeuTyMetGluAlaGlySerThrTrpAspPheValThrProGlnPro 340
QY 2050 CCACAACAGCTCTCTGCGCTTAATCAGTTGATCAGCTTTCCAAATCTGCATTTGTCTCTT 2109
DB 341 ProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu 360
QY 2110 TCTTCTTTGTAGCAAAACAATGCAGTTACGATCTCTCTACCAATCTCTCCAGCGCAAGAT 2169
DB 361 SerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAsp 380
QY 2170 TCTCATCTGCGAGTCATTGGTAGCACAACTGCTGCTTCTGTACAAATTAGTGGGCTATC 2229
DB 381 SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProile 400
QY 2230 TTTTGTAGGATTTGGATGATACAGCTTATCATAGTATGATTTGGCTAGGTTCTTAATCAA 2289
DB 401 PhePheGluAspLeuAspAspThrAlaTyArgAspTyArgTyLeuGlySerAsnGln 420
QY 2290 AAATCAATGCTCGAAATTAACAGTACGAGTACGAGTAAAGCCCAAGCTAATGCCCATCAGAT 2349
DB 421 LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAsp 440
QY 2350 TTGACTCTAGGGAATGAGATGCTAAGTATGCTATCAAGGAAGCTGGAAGCTT 2403
DB 441 LeuThrLeuGlyAsnGluMetProLysTyArgTyGlnGlnGlySerTyArgTyLeu 458

RESULT 33
US-10-766-711-36
; Sequence 36, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-36

Alignment Scores:
Pred. No.: 6,52e-200 Length: 458
Score: 2350.00 Matches: 458
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.8% Indels: 0
DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-766-711-36 (1-458)
```

```
QY 1030 GGGGAGCTTCTCAAGTAGTCACCAAGTTTCTCTGCTATGCTAACGAGGCTCCTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValThrSerPheSerAlaMetAlaAsnGluAlaProIleAla 20
QY 1090 TTTGTACGAATGTTGAGGAGTAAGAGGGGAGGAGTGTGCTGTTTCAGGATGGGCAG 1149
Db 21 PheValAlaAsnValAlaGlyValArgGlyGlyIleAlaValAlaGlnAspGlyGln 40
QY 1150 CAGGAGTGTATCATCTTCTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACT 1209
Db 41 GlnGlyValSerSerThrSerThrGluAspProValValSerPheSerArgAsnThr 60
QY 1210 GCGTAGAGTTTGTAGTGGAAAGTATGCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC 1269
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsn 80
QY 1270 GTTCTCTTCCGAAATAATGAAAAACCTGTTTCTCAACAATGTGCTTCTCTGTTTAC 1329
Db 81 ValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyr 100
QY 1330 ATTCTGCTAAGCAACCAAGTGGCAGGCTTCTTAATACGAGTAATAATTACGAGAT 1389
Db 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp 120
QY 1390 GGAGGAGCTATCTTCTGTAGAATGGTGGCAAGCAGGATCCAAATAACTCTGGATCAGTT 1449
Db 121 GlyGlyAlaIlePheCysLeuAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal 140
QY 1450 TCCTTTGTAGAGAGGAGTAGTTTCTTTTGTAGTACCAATGTAGTCTGGGAAAGGGGA 1509
Db 141 SerPheAspGlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyGly 160
QY 1510 GCTATTATGCGAAAAAGCTCCGTTGCTTAACGTGGCCCTGACAAATTTTAAAGGAAT 1569
Db 161 AlaIleTyrAlaLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn 180
QY 1570 ATCGCTAATGATGGTGGAGCAATTTATTTAGAGAACTTGGAGAGCTCAGTTTATCTGCT 1629
Db 181 IleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAla 200
QY 1630 GATTATGAGATATTTATTTTCATGGGAATCTTAAAGAACGCCAAAGAGATGCTGCC 1689
Db 201 AspTyrGlyAspIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAla 220
QY 1690 GATGTTAATGGGTAACTGTCTCTCACAGCCATTTGATGGGATCGGAGGGAATAA 1749
Db 221 AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIle 240
QY 1750 ACGACATTTAAGACTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGGCA 1809
Db 241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla 260
QY 1810 AACGAAATAACCAACCCAGCAGTCTTCCAAACTTCTAAATAATTAACGATGGTAAGA 1869
Db 261 AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGluGly 280
QY 1870 TACACAGGGATTTGTTTCTTAATGAACAGTACTTTCTACCAAAATGTTACGATA 1929
Db 281 TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle 300
QY 1930 GAGCAAGGAAGATTGTTCTTCTGTAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAG 1989
Db 301 GluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGln 320
QY 1990 ACAGTGGGAGTCTGTATATGAAAGCTGGAGTACATGGGATTTGTAACTCCACACCA 2049
Db 321 ThrGlyGlySerLeuTyrMetGluAlaGlySerThrTyrAspPheValThrProGlnPro 340
QY 2050 CCACACAGCTTCTGCGCTAATCAGTTGATCAGCTTCCAACTCCAACTGCAATTTGCTCTT 2109
Db 341 ProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu 360
QY 2110 TCTTCTTTGTAGCAAAACATGCAAGTTACGAATCTCTCTTACCAATCTCCAGCGCAAGAT 2169
```

```
Db 361 SerSerLeuLeuAlaAsnAsnAlaValThrAsnProThrAsnProThrAsnProAlaGlnAsp 380
QY 2170 TCTCATCTCTGCTAGTTCAGTTCAGCAACTGCTGTTCTGTTTACAATTTAGTGGGCTTATC 2229
Db 381 SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle 400
QY 2230 TTTTTCAGGATTTGGATGATACAGCTTATGATAGTATGATGGCTAGGCTTCTTAATCAA 2289
Db 401 PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGln 420
QY 2290 AAAATCAATCTCTGAAATTTACAGTTAGGACTTAAGCCCCAGCTAATGCCCATCAGAT 2349
Db 421 LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProAlaAsnAlaProSerAsp 440
QY 2350 TTGACTCTAGGAAATGAGATGCTTACCTAGTATGGCTATCAAGGAAGCTCGAAAGCTT 2403
Db 441 LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeu 458

RESULT 34
US-10-931-779-36
; Sequence 36, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: B104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-36
```

```
Alignment Scores:
Pred. No.: 6.52e-200 Length: 458
Score: 2350.00 Matches: 458
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.8% Indels: 0
DB: Gaps: 0

US-10-701-844-1 (1-4435) x US-10-931-779-36 (1-458)
```

```
QY 1030 GGGGAGCTTCTCAAGTAGTCACCAAGTTTCTCTGCTATGCTAACGAGGCTCCTATTGCC 1089
Db 1 GlyGlyAlaCysGlnValThrSerPheSerAlaMetAlaAsnGluAlaProIleAla 20
QY 1090 TTTGTACGAATGTTGAGGAGTAAGAGGGGAGGAGTGTGCTGTTTCAGGATGGGCAG 1149
Db 21 PheValAlaAsnValAlaGlyValArgGlyGlyIleAlaValAlaGlnAspGlyGln 40
QY 1150 CAGGAGTGTATCATCTTCTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACT 1209
Db 41 GlnGlyValSerSerThrSerThrGluAspProValValSerPheSerArgAsnThr 60
QY 1210 GCGTAGAGTTTGTAGTGGAAAGTATGCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC 1269
Db 61 AlaValGluPheAspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsn 80
QY 1270 GTTCTCTTCCGAAATAATGAAAAACCTGTTTCTCAACAATGTGCTTCTCTGTTTAC 1329
Db 81 ValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyr 100
QY 1330 ATTCTGCTAAGCAACCAAGTGGCAGGCTTCTTAATACGAGTAATAATTACGAGAT 1389
Db 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyrGlyAsp 120
```

QY	1390	GGAGAGCTATCTTCTGTAAGATCGTGGCAGCAGGATCCAAATAACTCTGGATCAGTT	1449
DB	121	GlyGlyAlaIlePheCysIysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerVal	140
QY	1450	TCCTTTGATGGAGAGGAGTGTCTTTCTTAGTACCAATGTAGCTGCTGGAAAGGGGA	1509
DB	141	SerPheAspGlyGlyValValPhePheSerSerAsnValAlaAlaGlySerGlyGly	160
QY	1510	GCTATTTATGCCAAAGCTCTCGGTGTCTAACTGGGCCCTGTACAAATTTTATAGGAAT	1569
DB	161	AlaIleTyrAlaIlysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsn	180
QY	1570	ATCGCTTAATGATGGCGGATTTATTTAGAGAGATCTGGAGAGCTCAGTTTATCTGCT	1629
DB	181	IleAlaAsnAspGlyAlaIleTyrLeuGlyGlySerGlyGlyLeuSerLeuSerAla	200
QY	1630	GATTATGGAGATATTATTTTCGATGGCAATCTTAAAGAACACAGCAGCAAGATGCTGCC	1689
DB	201	AspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAla	220
QY	1690	GATGTTAATGGCTAATCTGCTCTCAAGCCATTTTCGATGGGATCGGAGAGGAATA	1749
DB	221	AspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIle	240
QY	1750	ACGACATTAGAGCTAAAGCAGGCGATCAGATTCTCTTTAATGATCCCATCGAGATGCA	1809
DB	241	ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAla	260
QY	1810	AAAGGAAATAACAGCAGCGAGTCTTCCAAACTTCTAAATAATTAACGATGTGAAGGA	1869
DB	261	AsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLysIleAsnAspGlyGly	280
QY	1870	TACACAGGGGATATGTTTTGCTAATGGAAGCAGTACTTTGTACCAAAATGTTACGATA	1929
DB	281	TyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeuTyrGlnAsnValThrIle	300
QY	1930	GAGCAAGGAAGATGTTCTTCTGTAAGAGGCAAAATTTATCAGTGAATTTCTTAAGTCAG	1989
DB	301	GluGlnGlyArgIleValLeuArgGlyLysAlaLysLeuSerValAsnSerLeuSerGln	320
QY	1990	ACAGTGGGAGCTGTATATGGAAGCTGGGAGTACATGGATTTGTAACTCCACACCA	2049
DB	321	ThrGlyGlySerLeuTyrMetGluAlaGlySerThrThrPheValThrProGlnPro	340
QY	2050	CCACACAGCTCTCTGCGCTAATCAGTTGATCACGCTTTTCCAACTTCGCAATTTGCTCTT	2109
DB	341	ProGlnProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeu	360
QY	2110	TCTTCTTTGTAGCAAAATGAGTTACGATTCAGATTCCTTACCAATCTCCAGCGCAGAT	2169
DB	361	SerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAsp	380
QY	2170	TCTCATCTGCTCAGTCAATGCTAGCACAACTGCTGTTCTGTTCAATTTAGTGGCCTATC	2229
DB	381	SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle	400
QY	2230	TTTTTTTGGAGATTGGATGATACAGCTTATGATGGTATGATTTGGCTAGGTTCTTAATCAA	2289
DB	401	PhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGln	420
QY	2290	AAATCAATGCTCTGAAATTTACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCATGAT	2349
DB	421	LysIleAsnValLeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAsp	440
QY	2350	TTGACTCTAGGATGAGTCCCTAAGTATGGCTATCAAGGAGCTGGAAGCTT	2403
DB	441	LeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeu	458

RESULT 35

US-10-701-844-37

; Sequence 37, Application US/10701844

; Publication No. US20040067524A1

GENERAL INFORMATION:

; APPLICANT: Jackson, W. James
 ; APPLICANT: Pags, John
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
 ; FILE REFERENCE: 7969-086-999
 ; CURRENT APPLICATION NUMBER: US/10/701,844
 ; CURRENT FILING DATE: 2003-11-04
 ; PRIOR APPLICATION NUMBER: US/09/612,402B
 ; PRIOR FILING DATE: 2000-07-06
 ; PRIOR APPLICATION NUMBER: 08/942,596
 ; PRIOR FILING DATE: 1997-10-02
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 37
 ; LENGTH: 325
 ; TYPE: PRT
 ; ORGANISM: Chlamydia sp.
 ; US-10-701-844-37

Alignment Scores:
 Pred. No.: 37e-145 Length: 325
 Score: 1735.00 Matches: 325
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 22.0% Indels: 0
 DB: 4 Gaps: 0

US-10-701-844-1 (1-4435) x US-10-701-844-37 (1-325)

QY	2443	CTGAAAGCTACATGACATAAACTGGGTATAATCTCTGGCCCTGAGGAGTAGCTTCTTTG	2502
DB	1	LeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu	20
QY	2503	GTTCCAAATAGTTTATGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTC	2562
DB	21	ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisAlaIleGln	40
QY	2563	GCAAGTGTGGTGGGGCTCTTATTGTCTGAGAGATTATGGTTCTCGAGTTTCGAATTC	2622
DB	41	AlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhe	60
QY	2623	TTCTATCATGACCGCGATCTTTAGGTCAGGATATCGGTATATATAGTGGGGTTATTC	2682
DB	61	PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyTyrSer	80
QY	2683	TTAGGAGCAAACTCTTACTTTTGGATCATCGATGTTGGTCTAGCATTTACCGAGTATTT	2742
DB	81	LeuGlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe	100
QY	2743	GGTAGATCTAAAGATTATGTAGTGTCTGCTTCCAATCATCATCTTGCATAGGATCCGTT	2802
DB	101	GlyArgSerLysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerVal	120
QY	2803	TATCTATCTACCAACACAGCTTTATCTGGATCTTATTTGTCGGAGATGGTGGTTATCCGT	2862
DB	121	TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg	140
QY	2863	GCTAGCTAGCGTTTCGGGATTCAGCATATGAAACTCATATACATTTTCAGAGGAGAGC	2922
DB	141	AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer	160
QY	2923	GATGTTCTGGGATTAATCTCTGCTGGCTGGAGAGATTGGAGCGGATTTACCGATTGTG	2982
DB	161	AspValArgTrpAspAsnAsnCysLeuAlaGlyGlyIleGlyAlaGlyLeuProIleVal	180
QY	2983	ATTACTCCATCTAAGCTCTTATTTGAATGATTCGCTCTTCTTCGTGCAAGCTGAGTTTCT	3042
DB	181	IleThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSer	200
QY	3043	TATGCCGATCATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGGCATTCAGAGCGGA	3102
DB	201	TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly	220

```
QY 3103 CATCTCTAAATCTATCAGTTCCCTGTTGGAGTGAAGTTTGATCGATGTTCTTAGTACACAT 3162
Db 221 HisLeuLeuAanLeuSerValProValGlyValLysPheAspArgCysSerSerThrHis 240
QY 3163 CCTAATAATATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGT 3222
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260
QY 3223 ACTGAGACACGCTCTATCCATCCATCAAGACATGACACACAGATGCTCTTTTATTAGCA 3282
Db 261 ThrGluThrThrLeuLeuSerHisGlnGlnThrThrAspAlaPheHisLeuAla 280
QY 3283 AGACATGAGTTGGTTAGAGGATCTATGATGCTTCTTAACAGTAATATAGAAGTA 3342
Db 281 ArgHisGlyValValLargGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 300
QY 3343 TATGCCCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGGTTTGAGTGCAGGA 3402
Db 301 TyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGly 320
QY 3403 AGTAGAGTCCGGTTC 3417
Db 321 SerArgValArgPhe 325
RESULT 36
US-10-766-711-37
; Sequence 37, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; PRIORITY FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-37
Alignment Scores:
Pred. No.: 3 7e-145 Length: 325
Score: 1735.00 Matches: 325
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.0% Indels: 0
DB: 4 Gaps: 0
US-10-701-844-1 (1-4435) x US-10-766-711-37 (1-325)
QY 2443 CTGAAGCTACATGGAATAAACTGGGTATAATCTGGGCTGAGCGAGTAGCTTTTGG 2502
Db 1 LeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu 20
QY 2503 GTTCCAAATAGTTATGGGATCCATTTAGATATACGATCGCGCATTCAGCAATTCGA 2562
Db 21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40
QY 2563 GCAAGTGTGGTGGGCGCTTATGTCGAGGATATGGTTTCTCGAGTTTCGAAATTC 2622
Db 41 AlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhe 60
QY 2623 TTCTATCATGACCGCATGCTTTAGGTCAGGATATCGGTATATAGTGGGGTATATCC 2682
Db 61 PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyTyrSer 80
QY 2683 TTAGGAGCAACTCTTCTGATCATCGATGTTTGGTCTAGCATTTACCGAGTATTT 2742
```

```
Db 81 LeuGlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAGAAATATAGTAGTGTGCTGTTCCAAATCATCATCATAGTTCATAGGATCCGTT 2802
Db 101 GlyArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCCCAACAGCTTTATGTCGATCCCTATTGTTGCGGAGATGCGTTTATCCGT 2862
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg 140
QY 2863 GCTAGCTACGGGTTTGGGAATCAGCATATGAAACCTCATATACATATTTCAGAGAGAGAGC 2922
Db 141 AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer 160
QY 2923 GATGTTCTCGGATAAATACTGCTGCTGCGAGAGATTGGAGCGGATTTACCGATTGTG 2982
Db 161 AspValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180
QY 2983 ATTACTCCATCTAAGCTCTATTGTAATGAGTTGGTCTCTTTCGTGCAAGCTGAGTTTCT 3042
Db 181 IleThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATGCCGATCATGAATCTTTTACAGAGAGAGCGCATCAAGCTCGGGCATTTCAAGAGCGGA 3102
Db 201 TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly 220
QY 3103 CATCTCTAAATCTATCAGTTCCCTGTTGGAGTGAAGTTTGATGCGATGTTCTTAGTACACAT 3162
Db 221 HisLeuLeuAanLeuSerValProValGlyValLysPheAspArgCysSerSerThrHis 240
QY 3163 CCTAATAATATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGT 3222
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260
QY 3223 ACTGAGACACGCTCTATCCATCCATCAAGACATGACACACAGATGCTCTTTTATTAGCA 3282
Db 261 ThrGluThrThrLeuLeuSerHisGlnGlnThrThrAspAlaPheHisLeuAla 280
QY 3283 AGACATGAGTTGGTTAGAGGATCTATGATGCTTCTTAACAGTAATATAGAAGTA 3342
Db 281 ArgHisGlyValValLargGlySerMetTyrAlaSerLeuThrSerAsnIleGluVal 300
QY 3343 TATGCCCATGGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGGTTTGAGTGCAGGA 3402
Db 301 TyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGly 320
QY 3403 AGTAGAGTCCGGTTC 3417
Db 321 SerArgValArgPhe 325
RESULT 37
US-10-701-779-37
; Sequence 37, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-37
Alignment Scores:
Pred. No.: 3 7e-145 Length: 325
```

Score: 1735.00 Matches: 325
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.0% Indels: 0
DB: 5 Gaps: 0

US-10-701-844-1 (1-4435) X US-10-931-779-37 (1-325)

```
QY 2443 CTGAAGCTACATGAGTAACTCGGTATATCTCGGCGCTGAGGAGTACTCTTTG 2502
Db 1 LeuLysAlaThrTyrThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeu 20
QY 2503 GTTCCAAATAGTTATGGGATCCATTTAGATATACGATCTCGCATTCAGCAATCAA 2562
Db 21 ValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGln 40
QY 2563 GCAAGTGTGGGCGCTTATTGTCGAGGATATGCGTTCGAGTTCGAGTTTCGAATTC 2622
Db 41 AlaSerValAspGlyArgSerTyrCysArgLysLeuTrpValSerGlyValSerAsnPhe 60
QY 2623 TTCTATCATGACCGCATCTTTAGGTCAGGATATCGGTATATAGTGGGGTTATTC 2682
Db 61 PheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyTyrSer 80
QY 2683 TTAGGACCAACTCTTACTTTGGATCATCGATGTTGTTGTTAGCATTTACCGAATTT 2742
Db 81 LeuGlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAAGATATGATGTGCTGTTCCAAATCATCATCTTGCATAGGATCCGTT 2802
Db 101 GlyArgSerLysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCAACAAGCTTTATGTGATCTTATGTCGATCTTATGTCGAGATCGT 2862
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArg 140
QY 2863 GCTAGTACGGGTTTGGGATCAGCATATGAACCTCATATACATTTGTCAGAGGAGAGC 2922
Db 141 AlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSer 160
QY 2923 GATGTTCTGGGATATAACTCTGCTGGCTGGAGAGATTCGAGCGGATTTACCGATTGTG 2982
Db 161 AspValArgTrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleVal 180
QY 2983 ATTACTCCATCTAAGCTCTATTTGAATGAGTTCGCTCTTTCGTCGAAGCTGATTTCT 3042
Db 181 IleThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATGCCGATCATGAATCTTTTACAGGAAGGCGATCAAGCTCGGCATTCAGAGCGGA 3102
Db 201 TyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGly 220
QY 3103 CATCTCTAAATCTATCAGTTCCTGTTGGAGTGAAGTTTCATCGATGTTCTAGTACACAT 3162
Db 221 HisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerThrHis 240
QY 3163 CCTAATAAATATAGCTTTATGGGCGCTTATATCTGTGATGCTTATCCACCATCTCTGGT 3222
Db 241 ProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGly 260
QY 3223 ACTGAGACAACGCTCTATCCCATCAAGACATGGAACAGATGCTTTCATTTAGCA 3282
Db 261 ThrGluThrThrLeuLeuSerHisGlnGluThrTrpThrAspAlaPheHisLeuAla 280
QY 3283 AGACATCGAGTGTGGTTAGGAGTCTATGTTATCTCTTCTTAAACAAGTAATATAGAAGTA 3342
Db 281 ArgHisGlyValValValArgLysSerMetTyrAlaSerLeuThrSerAsnIleGluVal 300
QY 3343 TATGGCCATGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTGAGTCAGGA 3402
Db 301 TyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGly 320
QY 3403 AGTAGATCCGGTTC 3417
```

```
Db 321 SerArgValArgPhe 325
RESULT 38
US-09-841-132-325
; Sequence 325, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 325
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
; US-09-841-132-325
Alignment Scores:
Pred. No.: 3-86e-132 Length: 631
Score: 1590.50 Matches: 316
Percent Similarity: 88.7% Conservative: 7
Best Local Similarity: 86.8% Mismatches: 19
Query Match: 20.2% Indels: 22
DB: 4 Gaps: 4
US-10-701-844-1 (1-4435) X US-09-841-132-325 (1-631)
QY 3391 TTGAGTCGAGGAGTAGACTCGGTTCTAAATAATAT-----TGGTTAGATAGTTAA 3441
Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTrpArgPro----- 146
QY 3442 GTGTTAGCATGCTCTTTTCTTTGAGATCTACATCATTTTGTGTTTGTGTTTGTGTTGT 3501
Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164
QY 3502 TCCTATTCTGTATGATTCGCGAGCTCTCTCAAGTGTTAACGCTTAATGTAACCATCTCT 3561
Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrPro 184
QY 3562 TTTAAGGAGACATGTTTACTTTGAATGAGAGCTGCGCTTTTGTCAATGCTATGACGGA 3621
Db 185 PheLysGlyAspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGly 204
QY 3622 GCTGAAGAAGGTTTCGATTATCTCAGCTAATGGCGCAATTTAAGCATTAACGACAAAAC 3681
Db 205 AlaGluAsnGlySerIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsn 224
QY 3682 CATACATTATCATTTTACAGATTCTCAAGGCCAGTTCTTCAAAATATTATGCTTTCAATTC 3741
Db 225 HisThrLeuSerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSer 244
QY 3742 GCAGGAGACACTTACTCTGAGAGATTTTTCAGTCTGATGTTCTCGAAAAATGTTTCT 3801
Db 245 AlaGlyGluThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSer 264
QY 3802 TGGCGAGAAAGGAGATGATCTCCGGGAAACCGTGAGTATTTCCGAGAGCAGCGAAGTG 3861
Db 265 CysGlyGluLysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluVal 284
QY 3862 ATTTTCTGGGATAACTCCGTTGGGTATTTCTCTTTTATCTACTGTGCTGCAACCTCATATCA 3921
Db 285 IlePheTrpAspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrPro 304
QY 3922 ACTCCGCTGCT----- 3945
Db 305 ThrProProAlaProAlaProAlaProAlaSerSerSerLeuSerProThrValSer 324
```


3946	Qy	GATGCTCGGAAACGGTCTATTTTTCTGTAGAGACTAGTTGGAGACTCTCAGGCGTCAAA	4005
325	Db	AspAlaArgIySgIySerIlePheSerValGluThrSerLeuGluIleSerGlyValIyS	344
4006	Qy	AAAGGGTCATGTCGATATAATGCCGGNAATTCGGAACAGCTTTTCCAGGTAGAAT	4065
345	Db	LySgIyValMetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSer	364
4066	Qy	AATAATAATGCTGGTGGAGGCGAGTGGGTTCCGCTACACCATCAAGTACGACTTTTAC	4125
365	Db	AsnAsnAsnAlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheTh	382
4126	Qy	AGTTAAAAACTGTAAAGGAAAGTTTCTTTCCACAGATAAGTACGCTCTTCGCGAGCGG	4185
382	Db	rValIyAsnCylSgIySValSerPheThrAspAsnValAlaSerCylSgIyGlyI	402
4186	Qy	AGTGGTTTATAAGGCATTTGCTCTTTTCAAAGACAATGAAGGAGGCATATTTCTCCGAGG	4245
402	Db	yValValItyrIySgIyThrValLeuPheIySAspAsnGluGlyGlyIlePhePheArgG	422
4246	Qy	GAACACAGCATACCATGATTTAAAGGATTCCTGCTACTAATCATCAGGATCAGATAACGGA	4305
422	Db	yAsnThrAlaIyAspAspIeuGlyIleLeuAlaAlaThrSerArgGlnAsnThrGl	442
4306	Qy	GACAGGAGCGGTGGAGAGTATTGCTCTCTCCAGATGATTCTGTAAGTTTGAAGGCAA	4365
442	Db	uThrGlyGlyGlyGlyValIleCylSerProAspAspSerValIySPhesGluGlyAs	462
4366	Qy	TAAAGGTTCTATTGTTGATTACAATTTTGCAAAGCGCAGGCGGAAAGCATCTCTAAC	4425
462	Db	nIySgIySerIleValPheAspTyrAsnPheAlaIySgIyArgGlyGlySerIleLeuTh	482
4426	Qy	GAAGAAGATTC	4435
482	Db	rLysGluPhe	485

RESULT 39

```

US-10-872-155-325
/ Sequence 325, Application US/10872155
/ Publication No. US20040234536A1
/ GENERAL INFORMATION:
/ APPLICANT: Bhatia, Ajay
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Probst, Peter
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
/ TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
/ FILE REFERENCE: 210121.469C10
/ CURRENT APPLICATION NUMBER: US/10/872,155
/ CURRENT FILING DATE: 2004-06-18
/ PRIOR APPLICATION NUMBER: 09/620,412
/ PRIOR FILING DATE: 2000-07-20
/ PRIOR APPLICATION NUMBER: 09/598,419
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/556,877
/ PRIOR FILING DATE: 2000-04-19
/ PRIOR APPLICATION NUMBER: 09/454,684
/ PRIOR FILING DATE: 1999-12-03
/ PRIOR APPLICATION NUMBER: 09/426,571
/ PRIOR FILING DATE: 1999-10-22
/ PRIOR APPLICATION NUMBER: 09/410,568
/ PRIOR FILING DATE: 1999-10-01
/ PRIOR APPLICATION NUMBER: 09/288,594
/ PRIOR FILING DATE: 1999-04-08
/ PRIOR APPLICATION NUMBER: 09/208,277
/ PRIOR FILING DATE: 1998-12-08
/ NUMBER OF SEQ ID NOS: 599
/ SOFTWARE: FastSeq for Windows Version 3.0/4.0
/ SEQ ID NO 325
/ LENGTH: 631
/ TYPE: PRT
/ ORGANISM: Chlamydia trachomatis
US-10-872-155-325

```

Db 422 yAsnThrAlaTyAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrG1 442
 QY 4306 GACAGAGCGGTGGAGGATTTATTTCTCCAGATGTTCTGTAAGTTTGAAGGCNA 4365
 Db 442 uThrGlyGlyGlyGlyValIleCysSerProAspAspSerValIlePheGluGlyAs 462
 QY 4366 TAAAGGCTCTATTTGTTTATTAACAATTTTCAAAAGGAGGCGGAGGATCTTAAC 4425
 Db 462 nLysGlySerIleValPheAspTyAsnPheAlaLysGlyArgGlyGlySerIleLeuTh 482
 QY 4426 GAAAGATTC 4435
 Db 482 rLysGluPhe 485

RESULT 40
 US-09-841-260-95
 ; Sequence 95, Application US/09841260
 ; Publication No. US20030175700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Stromberg, Erika Jean
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
 ; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515
 ; CURRENT APPLICATION NUMBER: US/09/841,260
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 140
 ; SEQ ID NO 95
 ; LENGTH: 1016
 ; TYPE: PRT
 ; ORGANISM: Chlamydia trachomatis serovar D
 US-09-841-260-95

Alignment Scores:
 Pred. No.: 6,588-132 Length: 1016
 Score: 1589.00 Matches: 311
 Percent Similarity: 92.7% Conservative: 5
 Best Local Similarity: 91.2% Mismatches: 10
 Query Match: 20.2% Indels: 15
 DB: 3 Gaps: 2

US-10-701-844-1 (1-4435) x US-09-841-260-95 (1-1016)

QY 3451 ATGCCCTTTTCTTTGAGACTACATCATTTGTTTGTAGCTGTTGTTGTTCTATTCG 3510
 Db 1 MetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer 20
 QY 3511 TATGGATTCCGAGCTCTCTCAAGTGTAAAGCCTTAAGTAACTCCCTCTTTAAGGA 3570
 Db 21 TyrGlyPheAlaSerProGlnValLeuThrProAsnValThrThrProPheLysGly 40
 QY 3571 GACGATGTTTACTTGAATGAGAGTGGCTTTTGTCAATGCTATGAGGAGCTGAAGAA 3630
 Db 41 AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn 60
 QY 3631 GGTTCGATTATCTCAGCTAATGGCGACAATTTAAACGATTACCGGACAAAACCATACATTA 3690
 Db 61 GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu 80
 QY 3691 TCATTTACAGATTCTCAAGGCCAGTTCTTCAAAATTTATGCTTCATTTCCAGCAGGAGAG 3750
 Db 81 SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu 100
 QY 3751 ACATTTACTCTGAGAGATTTTTCGAGTCTGATGTTCTCGAAAATGTTTCTTTCGCGAGAA 3810
 Db 101 ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu 120
 QY 3811 AAGGGAATGATCTCCGGGAAAACCGTGAGTATTTCCGAGCAGCGGAAGTATTTCTGG 3870
 Db 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTyr 140
 QY 3871 GATAACTCCGTGGGGTATTTCTCTTTTATCTACTGTGCAACCTCATCATCACTCCGCT 3930

Db 141 AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro 160
 QY 3931 GCT-----CCACAGTTAGTATGATGCTCGG 3954
 Db 161 AlaProAlaProAlaProAlaSerSerLeuSerProThrValSerAspAlaArg 180
 QY 3955 AAAGGCTCTATTTTCTGTAGAGACTAGTTTGGAGATCTCAGCGCTCAAAAAGGGTTC 4014
 Db 181 LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValIleLysGlyVal 200
 QY 4015 ATGTTTCGATATAATATCCCGGAATTTCCGACAGTTCGAGGTAAAGATATAATAT 4074
 Db 201 MetPheAspAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAsn 220
 QY 4075 GCTGTGGTGGAGCGAGTGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAAAAA 4134
 Db 221 AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs 238
 QY 4135 CTGTAAGGGAAGTTTCTTTTCCAGATAACGCTAGCTCTTCGCGAGGCGGAGTGGTTTA 4194
 Db 238 nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyValValty 258
 QY 4195 TAAAGCATTGTGCTTTTCAAGAGCAATGAAGAGGCATATTTCTCCGAGGGAACACAGC 4254
 Db 258 rLysGlyThrValLeuPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAl 278
 QY 4255 ATACGATGATTTAAGGATTTCTGCTGCTACTAATCAGGATCAGAATACGAGACAGGAG 4314
 Db 278 aTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGly 298
 QY 4315 CGGTGGAGGAGTATTTGCTCTCCAGATGATTTCTGTAAAGTTTGAAGGCAATAAAGTTTC 4374
 Db 298 yGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySe 318
 QY 4375 TATTTGTTTGTATTACACTTTTGCAGGAGCGGAGCGGAGCATCTTACGAAAGATT 4434
 Db 318 rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh 338
 QY 4435 C 4435
 Db 338 e 338

Search completed: May 13, 2006, 12:11:20
 Job time : 896.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 13, 2006, 10:31:28 ; Search time 24.6 Seconds
(without alignments)
4471.544 Million cell updates/sec

Title: US-10-701-844-1

Perfect score: 7883

Sequence: 1 gggcaaaactctcccccg.....gcattcctaacgaagaattc 4435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-Q=/abes/ABSWEB_spool/US10701844/runat_12052006_165426_26373/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOFL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abs02p -USER=US10701844 -CGN 1 1 71 @runat_12052006_165426_26373
-NCPV=6 -ICPV=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WRAP TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5 COMB.pcp.*
2: /cgn2_6/prodata/1/1aa/6 COMB.pcp.*
3: /cgn2_6/prodata/1/1aa/H COMB.pcp.*
4: /cgn2_6/prodata/1/1aa/PCTUS COMB.pcp.*
5: /cgn2_6/prodata/1/1aa/RE COMB.pcp.*
6: /cgn2_6/prodata/1/1aa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	66.8	1012	2	US-09-612-402B-2
2	5267	66.8	1012	2	US-09-542-520-2
3	5116	64.9	984	2	US-09-612-402B-43
4	5106.5	64.8	1013	2	US-09-612-402B-15
5	5106.5	64.8	1013	2	US-09-612-402B-16
6	5106.5	64.8	1013	2	US-09-542-520-15
7	5106.5	64.8	1013	2	US-09-542-520-16
8	5090	64.6	1006	2	US-09-556-877-190
9	5090	64.6	1006	2	US-09-620-412C-190
10	5090	64.6	1006	2	US-09-598-419-190
11	5084	64.5	982	2	US-09-556-877-176
12	5084	64.5	982	2	US-09-620-412C-176

13	5084	64.5	982	2	US-09-598-419-176	Sequence 176, App
14	3336.5	42.3	670	2	US-10-197-220-169	Sequence 169, App
15	2552	32.4	505	2	US-09-612-402B-17	Sequence 17, Appl
16	2547	32.3	505	2	US-09-542-520-17	Sequence 17, Appl
17	2350	29.8	458	2	US-09-612-402B-36	Sequence 36, Appl
18	2350	29.8	458	2	US-09-542-520-36	Sequence 36, Appl
19	1735	22.0	325	2	US-09-612-402B-37	Sequence 37, Appl
20	1735	22.0	325	2	US-09-542-520-37	Sequence 37, Appl
21	1590.5	20.2	631	2	US-09-620-412C-325	Sequence 325, App
22	1590.5	20.2	631	2	US-09-598-419-325	Sequence 325, App
23	1589	20.2	1016	2	US-10-197-220-95	Sequence 95, Appl
24	1456	18.5	664	2	US-10-197-220-168	Sequence 168, App
25	1366.5	17.3	999	2	US-09-438-185A-455	Sequence 455, App
26	1359.5	17.2	973	2	US-09-430-723-2	Sequence 2, Appl
27	1129	14.3	928	2	US-09-428-132-2	Sequence 2, Appl
28	1094.5	13.9	967	2	US-09-438-185A-453	Sequence 453, App
29	1092	13.9	949	2	US-09-198-452A-478	Sequence 478, App
30	1083	13.7	930	2	US-09-198-452A-470	Sequence 470, App
31	1079	13.7	938	2	US-09-438-185A-448	Sequence 448, App
32	1041	13.2	947	2	US-09-438-185A-447	Sequence 447, App
33	1021	13.0	937	2	US-09-438-185A-449	Sequence 449, App
34	994.5	12.6	927	2	US-09-198-452A-472	Sequence 472, App
35	991.5	12.6	1414	2	US-09-438-185A-446	Sequence 446, App
36	981.5	12.5	780	2	US-09-438-185A-17	Sequence 17, Appl
37	973	12.3	932	2	US-09-438-185A-6	Sequence 6, Appl
38	964	12.2	922	2	US-09-198-452A-15	Sequence 15, Appl
39	943.5	12.0	866	2	US-09-438-185A-15	Sequence 15, Appl
40	828	10.5	1132	2	US-09-198-452A-466	Sequence 466, App
41	808	10.2	634	2	US-09-438-185A-451	Sequence 451, App
42	804	10.2	643	2	US-09-198-452A-474	Sequence 474, App
43	769.5	9.8	880	2	US-09-556-877-175	Sequence 175, App
44	769.5	9.8	880	2	US-09-620-412C-175	Sequence 175, App
45	769.5	9.8	880	2	US-09-598-419-175	Sequence 175, App

ALIGNMENTS

RESULT 1

US-09-612-402B-2
; Sequence 2, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OP INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-2

Alignment Scores:
Pred. No.: 0 Length: 1012
Score: 5267.00 Matches: 1012
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 66.8% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-612-402B-2 (1-1012)

QY 382 ATGCAGACCTCTTCCATGATCTTTCTTCAATGATTCAGCTTATCTTGCTGCTCT 441

Db 1 MetGlnThrSerPheHisLysPheLeuSerMetLeuAlaTyrSerCysSer 20

QY 442 TTAATGGGGGATATGAGCAGAAATCATGTTCTCTCAAGGAATTTACGATGGGAG 501
DB 21 LeuAenGlyGlyTyAlaAlaGluIleMetValProGlnGlyIleTyAspGlyGlu 40
QY 502 ACGTAACTGATCAATTTCCCTATACCTGTTATAGGAGATCGAGTGGACTACTGTTTTT 561
DB 41 ThrLeuThrValSerPheProTyThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGGTAAATTAATAAATCTTGACAAATCTTATGACGCTTTGCCCTTTAAGT 621
DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAATTTATAGGAGTTTTACTGTTTTTAGGAGAGACACTCGTTGACTTTC 681
DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGACTTCTACAAATGGGGAGCTCTAAGTAATAGCGTCTGTGATGGACTG 741
DB 101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120
QY 742 TTTTACTATTGAGGCTTTTAAGAAATATCTTTTCCAAATTCATTAATCTTGGCGTA 801
DB 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTTAATAAGGTAGCCAGACTCCGACGACAACTCTACACCGTCT 861
DB 141 LeuProAlaAlaThrAsnLysGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGGTACTATTATCTTAAACAGACTTTTGTGTTACTCAATATGAGAATCTCTATTC 921
DB 161 AsnGlyThrIleTySerLysThrAspLeuLeuLeuAsnGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAA 981
DB 181 TyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAAATTACAGCTTTGTGCTTCCAGAAATACTGCTCAAGCTGTAGTGGGGAGCTTGT 1041
DB 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyAlaCys 220
QY 1042 CAAGTAGTCAACAGTTTCTGCTATAGCTAACGAGGCTCTATTGCTTCTTTGAGCGAAT 1101
DB 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTCAGAGTAAGAGGGGGAGTCTGCTGTTTCAGGATGGCAGCAGGAGTGCA 1161
DB 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATPACTGCGGTAGAGTTT 1221
DB 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACGTAGCCGAGTAGGAGGAGGATTTACTCTACGGGAAGTTCCTTCTG 1281
DB 281 AspGlyAsnValAlaArgValGlyGlyIleTySerTyGlyAsnValAlaPheLeu 300
QY 1282 AATATGAAACCTTTCTTCTCAACATGTTGCTTCTCTGTTTACATGCTCCTAAG 1341
DB 301 AsnAsnGlyLysThrLeuPheLeuAsnAsnValAlaSerProValTyIleAlaAlaLys 320
QY 1342 CAACCAACAGTGGACAGCTTCTAATACGAGTAATAATTACGGAGATGGAGGACTATC 1401
DB 321 GlnProThrSerGlyGlnAlaSerAsnThrSerAsnAsnTyGlyAspGlyAlaIle 340
QY 1402 TTCTGTAGAATGGTGGCAGCAGAGATCCAAATACTCTGGATCAGTTCTCTTCATGGA 1461
DB 341 PheCysLysAsnGlyAlaGlnAlaGlySerAsnAsnSerGlySerValSerPheAspGly 360
QY 1462 GAGGAGTAGTTTCTTTAGTAGCAATGAGTCTGCGAAAGGGGAGCTATTTATGCC 1521
DB 361 GluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyAla 380
QY 1522 AAAAAGCTCTCGTCTAACTGTGGCCCTGTACAAATTTTTAAGGAATATCGCTAATGAT 1581

DB 381 LysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuArgAsnIleAlaAsnAsp 400
QY 1582 GGTGGAGCGATTTATTTAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTTAGGAGAT 1641
DB 401 GlyGlyAlaIleTyLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyGlyAsp 420
QY 1642 ATTATTTTCGATCGGAATCTTAAAGAACAGCAGCAAAGAGAAATGCTGCCGATGTTAATGGC 1701
DB 421 IleIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGly 440
QY 1702 GTAACGTGTCTCACAGCCATTTCCATGGGATCGGAGGGAATAAAGCAACATTAAGA 1761
DB 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460
QY 1762 GCTAAAGCAGGATCAGATTCTCTTAAATGATCCCATCGAGATGGCAACGGAATAAAC 1821
DB 461 AlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsn 480
QY 1822 CAGCCAGCGAGTCTTCCAAACTTCTAAATAATTAAACATCGTGAAGATACACAGGGAT 1881
DB 481 GlnProAlaGlnSerSerLysLeuLysIleAsnAspGlyGluGlyTyThrGlyAsp 500
QY 1882 ATTGTTTTCTTAATGGAACAGTACTTCTTACCAAAATCTTACGATAGACAGGAGG 1941
DB 501 IleValPheAlaAsnGlySerSerThrLeuTyGlnAsnValThrIleGluGlnGlyArg 520
QY 1942 ATTGTTCTTCTGTAAGGCAAAATATCAGTGAATTTCTTAAGTCACAGAGTGGAGT 2001
DB 521 IleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyLysSer 540
QY 2002 CTGTATATGGAAGCTGGAGTACATGGGATTTGTAACTCCACAAACCAACCAACAGCCT 2061
DB 541 LeuTyMetGluAlaGlySerThrTrpAspPheValThrProGlnProGlnGlnPro 560
QY 2062 CTTCCGCTAATCAGTTGATCAGCTTCCCAATCTGCAATTTGCTCTTCTCTTTGTTA 2121
DB 561 ProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeuLeu 580
QY 2122 GCAACCAATCAGTTACGATCTCTACCAATCTCCAGCGCAAGATTTCTCATCTGCA 2181
DB 581 AlaAsnAsnAlaValThrAsnProThrAsnProProAlaGlnAspSerHisProAla 600
QY 2182 GTCAATTGTAGCACAACTGCTGTTCTGTACAATTAGTGGGCTATCTTTTTGAGAT 2241
DB 601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAsp 620
QY 2242 TTGATGATACAGCTTATGATAGTATGATGGCTAGGTTCTTAATCAAAAATCAATCTC 2301
DB 621 LeuAspAspThrAlaTyAspArgTyAspTrpLeuGlySerAsnGlnLysIleAsnVal 640
QY 2302 CTGAAATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCATCAGATTTGACTCTAGG 2361
DB 641 LeuLysLeuGlnLeuGlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGly 660
QY 2362 AATGAGATGCTAAGTATGGCTATCAAGAAAGCTGGAAGCTTGGCTGGATCTTAATACA 2421
DB 661 AsnGluMetProLysTyGlyTyGlnGlySerTrpLysLeuAlaTrpAspProAsnThr 680
QY 2422 GCAATAATGTCCTTATACCTCTGAAGACTACATGACTAAACCTGAGCTTAATCTCGG 2481
DB 681 AlaAsnAsnGlyProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnProGly 700
QY 2482 CTGAGCAGTAGCTTCTTGTGTTCCAAATAGTTTATGGGATCCATCTTTAGATATACA 2541
DB 701 ProGluArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArg 720
QY 2542 TCTGCGCAATCAGCAATTCAGCAAGTGTGATGGGCGCTCTTATTTGTCAGGATATGCG 2601
DB 721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyCysArgGlyLeuTrp 740
QY 2602 GTTTCGAGGTTTCGAAATTTCTTCTATCATGACCGGATGCTTTAGGTCAAGGATATCGG 2661

741 ValSerGlyValSerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArg 760
2662 TATATTAGTGGGGTATATTCCTTAGGAGCAAACTCCTACTTTGGATCATCGATTTGGT 7721
761 TyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPheGly 780
2722 CTAGCAATTTACCGAAGTATTTGGTAGATCTAAAGATTATGTAGTGTGCTTCCCAATCAT 2781
781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHis 800
2782 CATGCTTGATAGGATCCGTTTATCTATCTACCCAAAGCTTTATGTGATCTCTATTGG 2841
801 HisAlaCysIleGlySerValTyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeu 820
2842 TTCGGAGATCGCTTATCCGTGCTAGCTACGAGGTGGGATCGGATATGAAACCTCA 2901
821 PheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSer 940
2902 TATACATTTGCAGAGAGCGATGTTCTGTTGGGATAAATCACTGCTGGCTGGAGAGATT 2961
841 TyrThrPheAlaGluGluSerAspValArgTrpAspAsnAsnCysLeuAlaGlyGluIle 860
2962 GGAGCGGGATTAACGATTTGATTAATCTCCATCTAAGCTCTTATTTGAATGAGTTGCGTCT 3021
861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrLeuAsnGluLeuA-gPro 880
3022 TTCGTCGAAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGAGGAGCGGATCAA 3081
881 PheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAspGln 900
3082 GCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCAGTCTCTGTTGGAGTGAAGTTT 3141
901 AlaArgAlaPheLysSerGlyHisLeuLeuAsnLeuSerValProValGlyValLysPhe 920
3142 GATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGGGCTTATCTGTGAT 3201
921 AspArgCysSerSerThrHisProAsnLysTyrSerPheMetAlaIleTyrIleCysAsp 940
3202 GCTTATCGCAACCATCTCTGTGATCTGAGACAACGCTCTCTATCCCATCAGAGCATGGACA 3261
941 AlaTyrArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThr 960
3262 ACAGATCGCTTTCATTTAGCAACATGAGTGTGTTAGAGATCTATGATGCTTCT 3321
961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrAlaSer 980
3322 CTAAAGTAATATAGAGTATATGCGCATGGAGATATCAGTATCGAGATGCTTCTCGA 3381
981 LeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArg 1000
3382 GGCTATGTTGATGCGAGGATGAGTCCGGTTC 3417
1001 GlyTyrGlyLeuSerAlaGlySerArgValArgPhe 1012

RESULT 2

US-09-542-520-2
; Sequence 2, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John L.
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia

US-09-542-520-2
Alignment Scores:
Pred. No.: 0 Length: 1012
Score: 5267.00 Matches: 1012
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 66.8% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-542-520-2 (1-1012)
QY 382 ATGCAAAAGCTCTTTCCATAAGTCTTTCTTCAATGATCTTAGCTTATTTCTGCTCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATATGGGGGGATATGCACAGAAATCATGGTTCCTCAAGGAATTTACGATGGGGAG 501
Db 21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 AGCTTAACCTGATCATTTCCCTTATCTGTATAGAGATCCGAGTGGGACTACTGCTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGAGAGATTAACTATAAAATCTTGACAATTTCTATTCAGCTTTGCTTTAAAGT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTGTTGGAACTTATAGGGAGTTTACTGTTTATAGGAGGAGGACACTGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGACATACGAGCTCTACAATGGGCGAGCTCTAAGTAATAGCCCTGCTGATGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTTAAAGAAATTTATCTTTTCCAATTTGCAATTTCACTTGCCTGA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAlaVal 140
QY 802 CTGCTCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACACATCTACACGCTCT 861
Db 141 LeuProAlaAlaThrAsnLysGlySerGlnThrProThrThrSerThrProSer 160
QY 862 AATGGTACTATTATTCTAAACAGATCTTTTCTACTCAATAATAGAGAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuLeuAsnGlnLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTTATAGATGCTAAGAGCTTTAACGGTTCAA 981
Db 181 TyrSerAsnLeuValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTCTTCCAAAGAAATACTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTCACAGTTTCTCTGCTATGGCTAAAGAGCTCTCTATGCTTTGTAGCGCAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTTCAGAGGACTAAGAGGGGAGGATTTGCTGCTTTCAGGATGGGACGAGGAGGTCTCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATATCTCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAACTAGCCCGAGTAGGAGGAGGATTTACTCTACCGGAGACGTTGCTTCTCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAANAACCTTGTGTTCTCTCAACAATGTTGCTTCTCTGTTTACATTGCTGCTAAG 1341

Db 301 AenAenGlyLeuThrLeuPheLeuAenAenValAlaSerProValTyrlleAlaAlaLys 320
Qy 1342 CAACCAACAGTGGACAGGCTTCTAATACGAGTAATAATACGGAGATGGAGGACTATC 1401
Db 321 GlnProThrSerGlyGlnAlaSerAenThrSerAenAenTyrlGlyAspGlyGlyAlaIle 340
Qy 1402 TTCTGTAAAGATGGGCCAAGCAGGATCCCAATAACTCTGGATCAGTTTCCTTGTATGGA 1461
Db 341 PheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAspGly 360
Qy 1462 GAGGAGTAGTTTCTTTAGTAGCAATGCTGCTGGCAAGGGAGCTATTATTATGCC 1521
Db 361 GluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyAlaIleTyrlAla 380
Qy 1522 AAAAGCTCTCGTTGTCTAACTGTGGCCCTGTACAAATTTTAAAGAAATATCGCTAATGAT 1581
Db 381 LysLysLeuSerValAlaAenCysGlyProValGlnPheLeuAenAenIleAlaAenAep 400
Qy 1582 GGTGAGCGATTTATTTAGGAGATCTCGAGAGCTCAGTTTATCTGCTGATTTATGGAGAT 1641
Db 401 GlyGlyAlaIleTyrlLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrlGlyAsp 420
Qy 1642 ATTTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAATGSC 1701
Db 421 IleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAenGly 440
Qy 1702 GTAACTGTCTCTCAAGACCATTTTCGATGGATCGGAGGAGAAATAACACATTAAGA 1761
Db 441 ValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArg 460
Qy 1762 GCTAAGCAGGCGATCAGATTTCTTTAATGATCCCATCGAGATGGCAAAACGGAATTAAC 1821
Db 461 AlaLysAlaGlyHisGlnIleLeuPheAenAenProIleGluMetAlaAenGlyAenAen 480
Qy 1822 CAGCCAGCGAGCTCTCCAACTCTTAAATTAACGATGGTGAAGGATACACAGGGAT 1881
Db 481 GlnProAlaGlnSerSerLysLeuLeuLysIleAenAenGlyGluGlyTyrlThrGlyAsp 500
Qy 1882 ATTGTTTTGTCTAAATGGAAGCATGTTTGTACCAAAATGTTACGATAGACAGGAAGG 1941
Db 501 IleValPheAlaAenGlySerSerThrLeuTyrlGlnAenValThrIleGluGlnGlyArg 520
Qy 1942 ATTGTTCTTGTGAAAGCAAAATATCAGTGAATTTCTTAAGTCAGACAGGTGGAGT 2001
Db 521 IleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGlnThrGlyGlySer 540
Qy 2002 CTGTATATGGAAGCTGGAGTACATGGGATTTTGTAACTCCCAACCAACCACCAAGCT 2061
Db 541 LeuTyrlMetGluAlaGlySerThrTrpAspPheValThrProGlnProGlnGlnPro 560
Qy 2062 CTGCGCGCTAAATCAGTTGATCAGCTTTTCCAACTCGCATTTGCTCTTTCTTTTGTGA 2121
Db 561 ProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeuSerSerLeuLeu 580
Qy 2122 GCAACAAATGAGTACGAATCTCTACCAATCTCCAGCGAAGATTTCTCATCTCGCA 2181
Db 581 AlaAenAenAlaValThrAenProProThrAenProProAlaGlnAenSerHisProAla 600
Qy 2182 GTCATTTGTAGCACACTGCTGGTTCTGTATCAATTAGTGGCCCTATCTTTTTCGAGGAT 2241
Db 601 ValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIlePhePheGluAen 620
Qy 2242 TTGGATGATACAGCTTATCATAGGTATGATGGCTAGGTTCTTAATTAACAAAATCAATGTC 2301
Db 621 LeuAenAenThrAlaTyrlAspArgTyrlAspTrpLeuGlySerAenGlnLysIleAenVal 640
Qy 2302 CTGAAATTAAGTTAGGAGCTAAGCCCCCAGCTAATATGCCCATCAGATTTGATCTTAGGG 2361
Db 641 LeuLysLeuGlnLeuGlyThrLysProProAlaAenAlaProSerAspLeuThrLeuGly 660
Qy 2362 AATGAGATGCTTAAGTATGCTATCAAGCAAGCTGGAGCTGGCGGATCCTTAATACA 2421

Db 661 AenGluMetProLysTyrlGlyTyrlGlnGlySerTrpLysLeuAlaIleTrpAspProAenThr 680
Qy 2422 GCAATTAATGCTCTTATACCTGAAAGCTACATGGAGCTAAACCTGGGTATAATCTCTGG 2481
Db 681 AlaAenAenGlyProTyrlThrLeuLysAlaIleThrTrpThrLysThrGlyTyrlAsnProGly 700
Qy 2482 CCTCAGCGAGTAGCTTCTTTGGTTCCAAATAGTTTATGGGATCCATTTAGATATACGA 2541
Db 701 ProGluArgValAlaSerLeuValProAenSerLeuTrpGlySerIleLeuAspIleArg 720
Qy 2542 TCTCGCATTCAGCAATTCAGCAAGTGTGGATGGCGCTCTTATTTGCGAGGATTTATGG 2601
Db 721 SerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrlCysArgGlyLeuTrp 740
Qy 2602 GTTTCGAGGTTTCGAATTTCTTATCATGACCGCATGCTTTAGTTCAGGTCAGGATATCGG 2661
Db 741 ValSerGlyValSerAenPheTyrlHisAspArgAspAlaLeuGlyGlnGlyTyrlArg 760
Qy 2662 TATATTAGTGGGGTTATCTCTTAGGAGCAACCTCTTCTTGGATCATCATGTTTGGT 2721
Db 761 TyrlSerGlyGlyTyrlSerLeuGlyAlaAenSerTyrlPheGlySerSerMetPheGly 780
Qy 2722 CTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATATTAGTAGTGTGCTGTTCCAATCAT 2781
Db 781 LeuAlaPheThrGluValPheGlyArgSerLysAspTyrlValValCysArgSerAenHis 800
Qy 2782 CATGCTTCATPAGATCCGTTTATCTATCTACCAACCAAGCTTTATGTGGATCTCTATTGG 2841
Db 801 HisAlaCysIleGlySerValTyrlLeuSerThrGlnGlnAlaLeuCysGlySerTyrlLeu 820
Qy 2842 TTCGAGATGGTTTATCCGCTAGCTACCGGTTTGGGAATCAGCATATCAAAACCTCA 2901
Db 821 PheGlyAspAlaPheIleArgAlaSerTyrlGlyPheGlyAenGlnHisMetLysThrSer 840
Qy 2902 TATCATTTGCAGAGGAGGAGCGATGTTCTGGTGGATATAAATCTGCTCGCTGGAGAGAT 2961
Db 841 TyrlThrPheAlaGluGluSerAspValArgTrpAenAenCysLeuAlaGlyGluIle 860
Qy 2962 GGAGCGGATTTACCGATTGTGATTTACTCCATCTAAGCTCTAATTTGAATGAGTTGGCTCT 3021
Db 861 GlyAlaGlyLeuProIleValIleThrProSerLysLeuTyrlLeuAenGluLeuArgPro 880
Qy 3022 TTGTGCAAGCTGAGTTTCTTATGCCGATCATGATCTTTTACAGAGAGAGCGCATCAA 3081
Db 881 PheValGlnAlaGluPheSerTyrlAlaAspHisGluSerPheThrGluGluGlyAspGln 900
Qy 3082 GCTCGGCGATTTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTGTTGGAGTGAAGTTT 3141
Db 901 AlaArgAlaPheLysSerGlyHisLeuLeuAenLeuSerValProValGlyValLysPhe 920
Qy 3142 GATCATGTTCTTAGTACACATCCTTAATAATATAGCTTTTATGGCGGCTTATATCTGTAT 3201
Db 921 AspArgCysSerSerThrHisProAenLysTyrlSerPheMetAlaAlaTyrlCysAsp 940
Qy 3202 GCTTATCCGACCATCTCTGGTACTGAGACACGCTCTATCCATCAAGAGACATGGACA 3261
Db 941 AlaTyrlArgThrIleSerGlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThr 960
Qy 3262 ACAGATGCTTTTCATTTAGCAAGACATGAGTGTGGTTAGAGGATCTATGTATGCTTCT 3321
Db 961 ThrAspAlaPheHisLeuAlaArgHisGlyValValValArgGlySerMetTyrlAlaSer 980
Qy 3322 CTAACAAATATATAGATATATGGCCATGGAAATATGATATGATATCGAGATCTCTTCTGA 3381
Db 981 LeuThrSerAenIleGluValTyrlGlyHisGlyArgTyrlGluTyrlArgAspAlaSerArg 1000
Qy 3382 GCCTATGCTTTTCAGTGCAGGAGTAGCTCCGGTTC 3417
Db 1001 GlyTyrlGlyLeuSerAlaGlySerArgValArgPhe 1012

RESULT 3

US-09-612-402B-43

; Sequence 43, Application US/09612402B

```
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-43

Alignment Scores:
Pred. No.: 0 Length: 984
Score: 5116.00 Matches: 984
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.9% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-612-402B-43 (1-984)
QY 466 GAAATCATGTTCTCTCAAGAAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTAT 525
DB 1 GluilewValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
QY 526 ACTGTTATAGGAGATCCGAGTGGGACTACTGTTTTTTCTGCAGGAGGTTAACTAAATA 585
DB 21 ThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLys 40
QY 586 AATCTTGACAAATCTATTGACGTTTGCCTTAAAGTTGTTTGGGACTATTATAGGAGT 645
DB 41 AsnLeuAspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySer 60
QY 646 TTTACTGTTTTAGGAGAGACACTCGTTCACCTTTTCGAGAACATACCGACTCTCAAAAT 705
DB 61 PheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsn 80
QY 706 GGGGCGAGCTTAAGTAATAGCGCTGCTGATGGACTGTTTACTATATGAGGTTTAAAGAA 765
DB 81 GlyAlaAlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGlu 100
QY 766 TTATCTTTTCCAAATTCATTTACTTCCGCTACTGCTGCTGCAACGACTAATAAG 825
DB 101 LeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLys 120
QY 826 GSTAGCCAGACTCCGAGACAAACATCTACACCGCTAATGCTACTATTTATTTCTAAACA 885
DB 121 GlySerGlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLysThr 140
QY 886 GATCTTTTGTACTCAATAATGAGAAGTTCTCATTTCTATAGTAATTTAGTCTCTGGAGAT 945
DB 141 AspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAsp 160
QY 946 GGGGGAGCTATAGATGTAGAGCTTAACGGTTCAAGGAATTACAGCTTTGTGCTTC 1005
DB 161 GlyGlyAlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPhe 180
QY 1006 CACAAAATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCACAGTTTCTCTGCT 1065
DB 181 GlnGluAsnThrAlaGlnAlaAspGlyCysGlnValThrSerPheSerAla 200
QY 1066 ATGCTTAACGAGGCTCTATTGCTTTGTAGCGAATGTTGAGGAGTAAGAGGGGGAGGG 1125
DB 201 MetAlaAsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGly 220
```

```
QY 1126 ATTGCTGCTGTTTCCAGATGGGAGGAGGAGTGTCTCATCTTCAACAGAGATCCA 1185
DB 221 IleAlaAlaValGlnAspGlyGlnGlyValSerSerSerThrThrGluAspPro 240
QY 1186 GTAGTAAGTTTTCCAGAAATCTCGGTAGAGTTTGTATGGGAACGTAGCCGAGTAGGA 1245
DB 241 ValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGly 260
QY 1246 GGAGGGATTTACTCTCTACGGGAACGTTGCTTCTCGAATAATGGAATAACCTTGTTC 1305
DB 261 GlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeu 280
QY 1306 AACAAATGTTGCTTCTCTGTTTACATTTGCTGTCTAAACCAACCAAGTCGACAGGTTCT 1365
DB 281 AsnAsnValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSer 300
QY 1366 AATACGAGTAATAATTACGGAGATGGAGGAGCTATCTTCTGTAAAGNATGGTCGCCAGCA 1425
DB 301 AsnThrSerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAla 320
QY 1426 GGATCCAATAACTCTGGATCAGTTTCTTGTGATGGAGGGAGTAGTATTTCTTTAGTAGC 1485
DB 321 GlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSer 340
QY 1486 AATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAGCTCTCGGTGCTAACTGT 1545
DB 341 AsnValAlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCys 360
QY 1546 GGCCCTGTACAATTTTAAAGGAATATCGCTAATGATGGTGGAGCGATTTATTAGAGAA 1605
DB 361 GlyProValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlu 380
QY 1606 TCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAA 1665
DB 381 SerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLys 400
QY 1666 AGAACACCAAGAGAGATCTGCCGATGTTAAATGGCGTAACTGTGCTCACAGGCATT 1725
DB 401 ArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIle 420
QY 1726 TCGATGGGATCGGGAGGGGAAATAACGACATTAAGAGCTTAACGACGGGCTACAGATCTC 1785
DB 421 SerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeu 440
QY 1786 TTTAATGATCCCATCCAGATGCGAACCGGAATAAACACGACGCGAGCTCTCCAAACTT 1845
DB 441 PheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeu 460
QY 1846 CTAATAAATTAACGATGCTGAAGGATACACAGGGGATATTGTTTTGCTAATGGAAGCAGT 1905
DB 461 LeuLysIleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSer 480
QY 1906 ACTTTGTACCAAAATGTTACGATAGACGAGGAGGATTTGTTCTTCTGTAAGAGGCAAAA 1965
DB 481 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 500
QY 1966 TTATCAGTGAATTTCTTAAGTCAGACGAGTGGGAGCTGTATATGGAAGCTGGAGTAGCA 2025
DB 501 LeuSerValAsnSerLeuSerGlnThrGlySerLeuTyrMetGluAlaGlySerThr 520
QY 2026 TGGGATTTTGTAATCTCACAAACCAACCAACGACGCTCTCCCGCTAATCAGTTGATCAGC 2085
DB 521 TrpAspPheValThrProGlnProProGlnGlnProProAlaAlaAsnGlnLeuIleThr 540
QY 2086 CTTTCCAATCTGCATTTGCTCTTCTTCTTGTGTAGCAAAACAATGAGTTACGAATCCT 2145
DB 541 LeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnPro 560
QY 2146 CCTACCAATCTCCAGGCGAAGATTTCTCATCTCGAGTCATTCGTAGCAACAACCTGCTGT 2205
DB 561 ProThrAsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGly 580
QY 2206 TCTGTTACAATTAGTGGGCTCTATCTTTTGGAGATTGGGATGATACAGCTTATGATAGG 2265
```

```
Db 501 SerValThrIleSerGlyProIlePheGluAspLeuAspPheAlaTyrAspArg 600
QY 2266 TATGATTGGCTAGGTTCTTAATCAAAAAATCAATGCTCCGAAATTACAGTTAGGGACTAAG 2325
Db 601 TyrAspTyrLeuGlySerAsnGlnIleAsnValLeuLeuGlnLeuGlyThrLys 620
QY 2326 CCCCAGACTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCCCTAAGTATGGCTAT 2385
Db 621 ProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyr 640
QY 2386 CAAGGAAGCTGAAGCTTCGGTGGGATCCTTAATACAGCAAAATAATGCTCTTATATCTGTG 2445
Db 641 GlnGlySerTyrPheLeuAlaTyrAspProAsnThrAlaAsnAsnGlyProTyrThrLeu 660
QY 2446 AAAGCTACATGACTGAACTGGGTATATCTCGGCTCGAGCGAGTAGCTTCTTTGGTT 2505
Db 661 LysAlaThrTyrThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuVal 680
QY 2506 CCAATATAGTTATGGGGATCCATTTTAGATATACGATCTGGCATTCAGCAATTCAGCA 2565
Db 681 ProAsnSerLeuTyrPheGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAla 700
QY 2566 AGTGTGGATGGCGCTCTTATTGTCGAGGATTATGGTTCTCGAGATTTCGAAATTTCTTC 2625
Db 701 SerValAspGlyArgSerTyrCysArgGlyLeuTyrPheValSerGlyValSerAsnPhePhe 720
QY 2626 TATCATGACCGGATGCTTTAGTTCAGGGATATCGGTATATTAGTGGGGTTATTCTTAA 2685
Db 721 TyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeu 740
QY 2686 GGAGCAAACTCTACTTTGGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATTGGT 2745
Db 741 GlyAlaAsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGly 760
QY 2746 AGATCTAAGATTATGATGTCGTTCCTCAATCATCATCGTTCGATAGGATCCGTTTAT 2805
Db 761 ArgSerLysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyr 780
QY 2806 CTATCTACCCAAACAAAGCTTTATGTGGATCTTATTTGTCGGAGATCGGTTATCCGTGT 2865
Db 781 LeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAla 800
QY 2866 AGCTACGGGTTGGGAATCAGCATATGAAACCTCATATACATTTTCGAGAGAGCGCAT 2925
Db 801 SerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluSerAsp 820
QY 2926 GTTCGTTGGGATAATAAAGTCTGCTGCTGGAGAGATTGGAGCGGATTACCGATTGTGATT 2985
Db 821 ValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIle 840
QY 2986 ACTCATCTAAGCTCTATTGTAATGAGTTGCGTCTCTTTCGTCGAAGCTGAGTTTCTTAT 3045
Db 841 ThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyr 860
QY 3046 GCCGATCATGAATCTTTTACAGAGAGCGGATCAAGCTCGGGCATTTCAAGACGGACAT 3105
Db 861 AlaAspHisGluSerPheThrGluGlyAspGlnAlaArgAlaPheLysSerGlyHis 880
QY 3106 CTCCTAAATCATAGTCTCTGTTGGAGTGAAGTTTGATCGATGTTCTAGTACACATCCT 3165
Db 881 LeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisPro 900
QY 3166 AATAAATATAGCTTTATGCGGCTTATATCTGTGATGCTTATCGCACCACTCTCTGTACT 3225
Db 901 AsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThr 920
QY 3226 GAGACAAACCTCTATCCCATCAAGAGACATGGAACACAGATGCCCTTCATTTAGCAGA 3285
Db 921 GluThrThrLeuLeuSerHisGlnGluThrTyrThrAspAlaPheHisLeuAlaArg 940
QY 3286 CATGGATTGTGGTAGAGATCTATGATGCTTCTTAAAGTAAATATAGAGTATAT 3345
```

```
Db 941 HisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyr 960
QY 3346 GGCCATGGAAGATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTACGTCAGGAAGT 3405
Db 961 GlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySer 980
QY 3406 AGAGTCCGGTTC 3417
Db 981 ArgValArgPhe 984
RESULT 4
US-09-612-402B-15
; Sequence 15, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-15
Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.4% Conservative: 17
Best Local Similarity: 96.7% Mismatches: 15
Query Match: 64.8% Indels: 1
DB: 2 Gaps: 1
US-10-701-844-1 (1-4435) x US-09-612-402B-15 (1-1013)
QY 382 ATGCAAAAGCTCTTCCATAAGTCTTCTTCAATGATTCTAGCTTATTCTTGCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaTyrSerCysSer 20
QY 442 TTAATGGGGGGGATATCGACGAAATCATGTTCTCTCAAGGAATTTACGATCGGGAG 501
Db 21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAATCTGATCATTTCCCTATCTGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGCAGAGTAAACATTAAAAATCTTCACAAATCTTCAGCAATTTCTTTCAGCTTTTAACT 621
Db 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
QY 622 TGTGTTGGGAACTTATTAGGGAGTTTACTGTGTTTGGGAGAGGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACCGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTTTACTATTGAGGGTTTAAAGAAATTATCTTTTCCAAATTGCAATTCATTACTTGCCTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuAlaVal 140
QY 802 CTGCTGCTGCAACCACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACACGCTCT 861
Db 141 LeuProAlaAlaThrThrAsnAsnGlySerGlnThrProSerThrThrSerThrProSer 160
```


Db 881 ProPheValGlnAlaGluPheSerTyrAlaAspHisGluSerPheThrGluGluGlyAsp 900
QY 3079 CAAGCTCGGCATTCACAGCGGACATCTCTAAATCTATCATGTTCTCTGTTGGAGTGAAG 3138
Db 901 GlnAlaArgAlaPheArgSerGlyHisLeuMetAenLeuSerValProValGlyVallys 920
QY 3139 TTGATCATGTTCTAGTACACATCCTAATAATATAGCTTTATGGCGGCTTATCTGT 3198
Db 921 PheAspArgCysSerThrHisProAenIlySerPheMetGlyAlaIyriIeCys 940
QY 3199 GATGCTTATCGCACCATCTCTGGTACTCGACACAGCTCTCCATCCATCAAGAGACATGG 3258
Db 941 AspAlaIyriArgThrIleSerGlyThrGlnThrLeuLeuSerHisGlnGluThrTrp 960
QY 3259 ACAACAGATGCTTTTCATTTAGCAAGACATGAGTGTGGTTAGAGGATCTATGATGCT 3318
Db 961 ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
QY 3319 TCTCTAACAGTAATATATAGATATATGCGCATGGAATATAGATATGAGATGCGAGATCTTCT 3378
Db 981 SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
QY 3379 CGAGCTATGTTTTCAGTGCAGGAAGTAGAGTCCGGTTC 3417
Db 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 5

US-09-612-402B-16
; Sequence 16, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-16

Alignment Scores:
Pred. No.: 0 Length: 1013
Score: 5106.50 Matches: 980
Percent Similarity: 98.2% Conservative: 15
Best Local Similarity: 96.7% Mismatches: 17
Query Match: 64.8% Indels: 1
Gaps: 2

US-10-701-844-1 (1-4435) x US-09-612-402B-16 (1-1013)

QY 382 ATGCAACGCTCTTCCATTAAGTTCTTCTTCAATGATTTCTAGCTTATTTCTGTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPheLeuSerMetIleLeuAlaIyriSerCysSer 20
QY 442 TTAATATGGGGGATATGACAGCAAAATCATGTTCTCTCAAGGAATTTACATGGGAG 501
Db 21 LeuThrGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAATGATATCTTCCCTATATGTTATAGAGATCCGAGTGGGACTACTGTTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCGAGGAGGTAAACATTAATAAATCTTGACAAATCTTATGTCAGCTTTGCTTTAAGT 621

Db 61 SerAlaGlyGluLeuThrLeuLysAenLeuAspAenSerIleAlaAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAACCTTATTAGGAGTTTATCTGTTTTAGGAGAGACACTCTGTTGACTTTC 681
Db 81 CysPheGlyAenLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGAACATACGAGCTTCTACAAATGGGGAGCTCTAAGTAATAGCGCTGCTGATGGACTG 741
Db 101 GluAenIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAenSerGlyLeu 120
QY 742 TTTTACTATGAGGCTTTTAAAGAAATATCTCTTTTCCAAATTCAAATTCATTACTTCCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAenCysAenSerLeuLeuAlaVal 140
QY 802 CTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCGACGACAACTCTACACGCTCT 861
Db 141 LeuProAlaAlaThrThrAenAenGlySerGlnThrProThrThrThrSerThrProSer 160
QY 862 AATGGTACTATTTTATTCTAAACAGATCTTTTGTACTCAATAATAGAAAGTCTCATTC 921
Db 161 AenGlyThrIleTyrSerLysThrAspLeuLeuLeuAenAenGluLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAAGAGCTTAACGGTTCAA 981
Db 181 TyrSerAenLeuValSerGlyAspGlyGlyThrIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTGTTCTTCCAGAAATAATCTGCTCAAGCTGATGGGGAGCTTGT 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAenThrAlaGlnAlaAenGlyGlyAlaCys 220
QY 1042 CAAGTACTCACAGTCTTCTGCTATGCTTAACGAGCTCTTATTCCTTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAlaPheIleAlaAen 240
QY 1102 GTTCAGGAGTAACAGGGGAGGAGTCTGCTCTTCAAGTATGGGACGACAGGAGTGTCA 1161
Db 241 ValAlaGlyValArgGlyGlyIleAlaAlaValGlnAenGlyGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAGTTTTTCCAGAAATATATCTGCGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAenThrAlaValGluPhe 280
QY 1222 GATGGAGCTAGCCCGAGTAGGAGGAGGATTTACTCTCTACGGGAAGCTTCTTCTCTG 1281
Db 281 AspGlyAenValAlaArgValGlyGlyIleTyrSerTyrGlyAenValAlaPheLeu 300
QY 1282 AATAATGAAAAACCTTGTCTTCTCAAAATGTTGCTTCTCTCTGTTTCAATTGCTGCTAAG 1341
Db 301 AenAenGlyLysThrLeuPheLeuAenAenValAlaSerProValTyrIleAlaAlaGlu 320
QY 1342 CAACCAACAGTGCAGGCTTCTTAATACGAGTAATAATACGAGTAGGAGGAGCTATTC 1401
Db 321 GlnProThrAenGlyGlnAlaSerAenThrSerAspAenTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTAGAATGCTGGCCAA--GCAGATCCCAATCTCTGATCTGATCTTCTTCTGAT 1458
Db 341 PheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerValSerPheAsp 360
QY 1459 GGAGAGGAGTAGTTTTCTTTAGTAGCAATGTAGCTCTGGGAAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGlyAlaIleTyr 380
QY 1519 GCCAAAAAGCTCTCGTTGCTAATCTGTGGCCCTGTACAAATTTTAAAGAAATATCCCTAAT 1578
Db 381 AlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuGlyAenIleAlaAen 400
QY 1579 GATGTTGAGCGATTTATTTAGGAATCTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTTTTCGATGGGAATCTTAAAGAAACAGCAAGAAATGCTGCCGATGTTAAT 1698
Db 421 AspIleIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAlaAspValAen 440

1699 GCGTAACTGCTCTCACAAGCCATTTCGATGGATCGGAGGGAATAATACGACATTA 1758
 Db GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyIleThrThrLeu 460
 1759 AGAGCTAAAGCAGGGGATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
 Db ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAen 480
 1819 AACAGCCAGCCAGCTCTTCCAAACTTCTAAAAATTAAAGATGGTGAAGGATACACAGG 1878
 Db AsnGlnProAlaGlnSerSerGlnProLeuLysIleAsnAspGlyGluGlyThrGly 500
 1879 GATATTGTTTCTTAATCGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGCA 1938
 Db AspIleValPheAlaAsnGlyAsnSerThrLeuLysGlnAsnValThrIleGluGlnGly 520
 1939 AGGATTGTTCTTCGTGAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGGTGGG 1998
 Db ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
 1999 AGTCTGTATATGAAGCTGGAGTACATGGGATTTGTAACTTGTAACTCCACACACACAG 2058
 Db SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProGlnGln 560
 2059 CCTCTGCGCCTAATCAGTTGATCAGCTTCCAACTCGCATTTGTCTTCTTCTCTTGG 2118
 Db ProProAlaAlaAsnGlnLeuIleThrLeuSerAsnLeuHisLeuSerLeuSerLeu 580
 2119 TTAGCAAAACAATGACGATACGATCTCTTACCAATCTCCAGCGCAAGATTTCTCATCCT 2178
 Db LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
 2179 GCAGTCATTGGTAGCAACTGCTGTTCTGTACAAATTAGTAGGGGCTATCTTTTGGAG 2238
 Db AlaValIleGlySerThrAlaGlyProValThrIleSerGlyProPhePheGlu 620
 2239 GATTTCGATGATACAGCTTATGATAGTATGATGGTCTAGTCTTCTAATCAAAAAACAAT 2298
 Db AspLeuAspAspThrAlaTyrAspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsp 640
 2299 GTCTCAAAATTAAGTATGAGGACTAAGCCCAAGCTAATGCCCAATCAGATTTGACTCTA 2358
 Db ValLeuLysLeuGlnLeuGlyThrGlnProSerAlaAsnAlaProSerAspLeuThrLeu 660
 2359 GGGAAATGAGATGCTTAAGTATGCTATCAAGAGCTGGAAGCTTGGCTGGGATCTCTAAT 2418
 Db GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrpLysLeuAlaTrpAspProAsn 680
 2419 ACAGCAAAATTAAGTCTTATCTCTGAAAGCTACATGGACTAAAGCTGGTATATCTCT 2478
 Db ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnPro 700
 2479 GGGCTGAGGAGTGTCTCTTGGTTCCTCAATAGTTTATGGGATCCATTTTAGATATA 2538
 Db GlyProGlnArgValAlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIle 720
 2539 CGATCTGCGCATTCAGCAATTCAGCAAGTGTGGATGGGCTCTTATTCGCGAGGATTA 2598
 Db ArgSerAlaHisSerAlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeu 740
 2599 TGGTTTCTGAGTTCGAATTTCTCTATCATGACCGCATGCTTTAGCTCAGGATAT 2658
 Db TrpValSerGlyValSerAsnPheSerTyrHisAspArgAspAlaLeuGlyGlnGlyTyr 760
 2659 CGGTATATTAGTGGGGTTATCTCTAGGAGCAAACTCTTACTTTGGATCATCGATGTTT 2718
 Db ArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsnSerTyrPheGlySerSerMetPhe 780
 2719 GGTCTAGCATTTACCGAAGTATTTGGTAGATCTAAAGATTATGATGATGCTGCTGCTCAAT 2778
 Db GlyLeuAlaPheThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsn 800

2779 CATCATGCTGCATAGGATCCGTTTATCTATCTATCTATCCCAACAGCTTTATGCGATCCTAT 2838
 Db HisHisAlaCysIleGlySerValTyrLeuSerThrLysGlnAlaLeuCysGlySerTyr 820
 2839 TTGTTCCGAGATCGCTTTTATCCGTCTAGCTACGGGTTTGGGAATCAGCATAGAAACC 2898
 Db LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThr 840
 2899 TCATATACATTTGCAGAGGAGGAGCATGTTCTGTTGGGATAATACTGTCTGCTGAGAG 2958
 Db SerTyrThrPheAlaGluSerAspValArgTrpAspAsnAsnCysLeuValGlyGlu 860
 2959 ATTGAGCGGAGATTACGATTTGATCTCATCTCACTTAACTCTTATTTGAATGAGTTGCT 3018
 Db IleGlyValGlyLeuProIleValThrThrProSerLysLeuTyrLeuAsnGluLeuArg 880
 3019 CTTTCTGTCAGCTGAGTTTCTTATCCGATCATGAATCTTTTACAGAGGAAGCGAT 3078
 Db ProPheValGlnAlaGluPheSerTyrAlaAspHisGlySerPheThrGluGluGlyAsp 900
 3079 CAAGCTCGGGCATTCAAGAGCGGACATCTCTAAATCTATCAGTTCTCTTGGAGTGAAG 3138
 Db GlnAlaArgAlaPheArgSerGlyHisLeuMetAsnLeuSerValProValGlyValLys 920
 3139 TTTGATCGATGTTCTAGTACACATCTCTAATAATATAGCTTTATGCGGCTTATATCTGT 3198
 Db PheAspArgCysSerSerThrHisProAsnLysTyrSerPheMetGlyAlaTyrIleCys 940
 3199 GATGCTTATCGCACCATCTCTGCTACTGAGACACGCTCTCTATCCCATCAAGAGACATGG 3258
 Db AspAlaTyrArgThrIleSerGlyThrGlnThrThrLeuLeuSerHisGlnGluThrTrp 960
 3259 ACAACAGATGCTCTTCATTAGCAAGACATGCGATTGTCTGTAGAGGATCTATGTATGCT 3318
 Db ThrThrAspAlaPheHisLeuAlaArgHisGlyValIleValArgGlySerMetTyrAla 980
 3319 TCTTAACAGTAATATAGAACTATATAGCCATGGAAGATATGAGTATCGAGATGCTTCT 3378
 Db SerLeuThrSerAsnIleGluValTyrGlyHisGlyArgTyrGluTyrArgAspThrSer 1000
 3379 CAGGCTATGTTGAGTCAGGAGTAGAGTCCGGTTC 3417
 Db ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013

RESULT 6

US-09-542-520-15
 ; Sequence 15, Application US/09542520
 ; Patent No. 6887843
 ; GENERAL INFORMATION:
 ; APPLICANT: Jackson, W. James
 ; APPLICANT: Pace, John L.
 ; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
 ; FILE REFERENCE: 7969-076-999
 ; CURRENT APPLICATION NUMBER: US/09/542,520
 ; CURRENT FILING DATE: 2000-04-03
 ; PRIOR APPLICATION NUMBER: PCT/US98/20737
 ; PRIOR FILING DATE: 1998-10-01
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 1013
 ; TYPE: PRT
 ; ORGANISM: Chlamydia
 US-09-542-520-15

Alignment Scores:
 Pred. No.: 0 Length: 1013
 Score: 5106.50 Matches: 980
 Percent Similarity: 98.4% Conservative: 17
 Best Local Similarity: 96.7% Mismatches: 15
 Query Match: 64.8% Indels: 1
 DB: 2 Gaps: 1

US-10-701-844-1 (1-4435) x US-09-542-520-15 (1-1013)

QY 382 ATGCAACGCTCTTCCATAAGTTCTTTCTTTCAATGATCTAGCTTATTTCTGCTCT 441
Db 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleLeuAlaTyrSerCysCysSer 20
QY 442 TTAATGGGGGGGATATGCAGCAGAAATCATGGTTCTTCAAGGAATTCAGTGGGAG 501
Db 21 LeuAsnGlyGlyTyrAlaAlaGluIleMetValProGlnGlyIleTyrAspGlyGlu 40
QY 502 ACGTTAACTGATCATTTCCCTATATCTGTATAGCAGATCCGAGGCGACTCTTTT 561
Db 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
QY 562 TCTGCAGGAGAGTTAAACATTAATAAATCTTGCAATTTCTATTGCAGCTTTGCTTTAAGT 621
Db 61 SerAlaGlyGlyLeuThrLeuLysLeuLeuAspAsnSerIleAlaLeuProLeuSer 80
QY 622 TGTTTTGGGAACCTTATTAGGAGTTTACTGTGTTTGGGAGGAGACACTCGTTGACTTTC 681
Db 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
QY 682 GAGACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGACTG 741
Db 101 GluAsnIleArgThrSerThrAsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeu 120
QY 742 TTACTATTAGGGTTTTAAAGAATTTATCTTTTCCAAATGCAATTCATTACTTCCCGTA 801
Db 121 PheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnProLeuLeuAlaVal 140
QY 802 CTGCTGCTGCACGACTTAATAAGGGTAGCGAGCTCGACGACGCAACATCTACACGGTCT 861
Db 141 LeuProAlaLeuThrThrAsnAsnGlySerGlnThrProSerThrSerThrProSer 160
QY 862 ATGTGACTATTATTCTAAACAGATCTTTTGTACTCAATAATGAGAGTTCTCATTC 921
Db 161 AsnGlyThrIleTyrSerLysThrAspLeuLeuLeuAsnAsnGlnLysPheSerPhe 180
QY 922 TATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTTAAGAGCTTAAACGGTTCAA 981
Db 181 TyrSerAsnSerValSerGlyAspGlyGlyAlaIleAspAlaLysSerLeuThrValGln 200
QY 982 GGAATTAGCAAGCTTTGTCTTCCAGAAATATCTGCTCAAGCTGATGGGGAGCTGTG 1041
Db 201 GlyIleSerLysLeuCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCys 220
QY 1042 CAAGTAGTCACCAAGTTCTCTGCTATGGCTAACGAGGCTCTATTGCTTGTAGCGAAT 1101
Db 221 GlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheValAlaAsn 240
QY 1102 GTTGAGGAGTAAGAGGGGAGGATTTGCTGTTTCAGATGGGCGAGGAGGTGCA 1161
Db 241 ValAlaGlyValArgGlyGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSer 260
QY 1162 TCATCTACTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATATCTGGGTAGAGTTT 1221
Db 261 SerSerThrSerThrGluAspProValValSerPheSerArgAsnThrAlaValGluPhe 280
QY 1222 GATGGGAAGTACCGGAGTAGGAGGATTTACTCTACGAGGAGCTTCTTCTTCCTG 1281
Db 281 AspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeu 300
QY 1282 AATAATGGAAAAACCTTGTCTCAACAAATGTTGCTTCTCTCTGTTTACATGCTGCTAAG 1341
Db 301 AsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAlaGlu 320
QY 1342 CAAACCAAGTGGAGCGCTTCTAATACAGTAATAATACGAGATCGGAGAGCTATC 1401
Db 321 GlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIle 340
QY 1402 TTCTGTAAAGTGGTGGCAA--GCAGGATCCAATAACTCTGGATCAGTTTCTTTGAT 1458
Db 341 PheCysLysAsnGlyAlaGlnAlaIaGlySerAsnAsnSerGlySerValSerPheAsp 360

QY 1459 GGAGGAGAGTAGTTTCTTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGAGCTATTTAT 1518
Db 361 GlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAlaIleTyr 380
QY 1519 GCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAAATTTTAAAGCAATATCCCTAAT 1578
Db 381 AlalysLysLeuSerValAlaAsnCysGlyProValGlnLeuLeuGlyAsnIleAlaAsn 400
QY 1579 GATGTCGAGCGCATTTATTAGGAGATCTCGAGAGCTCAGTTTATCTGCTGATTATGGA 1638
Db 401 AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGly 420
QY 1639 GATATTATTTTCGATGGGAATCTTAAAGAACAGCAAGAGATGCTGCGCATTTAAT 1698
Db 421 AspMetIlePheAspGlyAsnLeuLysArgThrAlaLysGluAsnAlaAlaAspValAsn 440
QY 1699 GGCCTAACGTGCTCTCACAGCCATTTCCGATCGGATCGGAGGGGAAATAACGACATTA 1758
Db 441 GlyValThrValSerSerGlnAlaIleSerMetGlySerGlyGlyLysIleThrThrLeu 460
QY 1759 AGAGCTAAAGCAGGCGCATCAGATTCTCTTTAATGATCCCATCGAGATGGCAACGAAAT 1818
Db 461 ArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspProIleGluMetAlaAsnGlyAsn 480
QY 1819 AACGACGAGCGCGAGCTTCCAAACTTCTAAAAATTAACGATGGTGAAGGATACACAGG 1878
Db 481 AsnGlnProAlaGlnSerSerGluProLeuLysIleAsnAspGlyGluGlyTyrThrGly 500
QY 1879 GATATTGTTTTGCTTAATGGAAGCAGTACTTTGTACCAAAATGTTACCATAGACGACAGA 1938
Db 501 AspIleValPheAlaAsnGlyAsnSerThrLeuTyrGlnAsnValThrIleGlnGly 520
QY 1939 AGGATTGTTCTTCGTGAAAAGGCAAAATTTATCAGTGAATTTCTTAAGTCAGACAGTGGG 1998
Db 521 ArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeuSerGlnThrGlyGly 540
QY 1999 AGTCGTATATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACAAACCACACACAG 2058
Db 541 SerLeuTyrMetGluAlaGlySerThrLeuAspPheValThrProGlnProProGlnGln 560
QY 2059 CCTCCGCGCTTAATCAGTTGATCAGCTTCCCATCTGCAATCTGCTCTCTTCTCTCTTCTTG 2118
Db 561 ProProAlaAlaAsnGlnSerIleThrLeuSerAsnLeuHisLeuSerLeuSerSerLeu 580
QY 2119 TTAGCAAAACAATGCAAGTTACGAATCTCTACCAATCTCTCCAGCGCAAGATTTCTCATCT 2178
Db 581 LeuAlaAsnAsnAlaValThrAsnProProThrAsnProProAlaGlnAspSerHisPro 600
QY 2179 GCAGTCATTTGGTAGCAACCTGCTGGTCTGTATCAATTAGTGGGCCCTATCTTTTGTAG 2238
Db 601 AlaValIleGlySerThrAlaGlySerValThrIleSerGlyProIlePhePheGlu 620
QY 2239 GATTTGGATGATACAGCTTATGATAGGTATGATGCTAGTTCCTTAATCAAAAAATCAAT 2298
Db 621 AspLeuAspAspThrAlaTyrAspArgTyrAspTyrLeuGlySerAsnGlnLysIleAsp 640
QY 2299 GTCTCTGAAATACAGTTAGGGACTTAAGCCCCCAGCTTAATGCCCCCATCAGATTTGACTCTA 2358
Db 641 ValLeuLysLeuGlnLeuGlyThrGlnProProAlaAsnAlaProSerAspLeuThrLeu 660
QY 2359 GGGAAATGATGCTCTTAAGTATGGCTATCAAGGAAGCTGGAAGCTGCGTGGGATCTTAAT 2418
Db 661 GlyAsnGluMetProLysTyrGlyTyrGlnGlySerTyrLysLeuAlaTyrAspProAsn 680
QY 2419 ACAGCAATAATGCTCTTACTCTGAAGCTACATGAGCTTAAAGCTGGGTATATCTCT 2478
Db 681 ThrAlaAsnAsnGlyProTyrThrLeuLysAlaThrTyrThrLysThrGlyTyrAsnPro 700
QY 2479 GGGCCTGAGCGAGTAGTCTTTTGTGTTCCAAATAGTTTATGGGAGTCCATTTTAGATATA 2538
Db 701 GlyProGluArgValAlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIle 720


```
||||| 1001 ArgGlyTyrGlyLeuSerAlaGlySerLysValArgPhe 1013
Db
RESULT 8
US-09-556-877-190
; Sequence 190, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Peter
; APPLICANT: Probat, Peter
; APPLICANT: Bhacia, Ajay
; APPLICANT: Skeiky, Yaeir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469CS
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-190

Alignment Scores:
Pred. No.: 0 Length: 1006
Score: 5090.00 Matches: 979
Percent Similarity: 99.9% Conservative: 2
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 64.6% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-556-877-190 (1-1006)
QY 472 ATGGTTCTCAAGGAATTTACGATGGGAGAGCTTAACGTGATCATCTTCCCTATACGTGT 531
Db 25 MetIleProGInGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44
QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCTCGCAGGAGAGTTAACATTAATAAATCTT 591
Db 45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64
QY 592 GACAACTTATTCGAGCTTGGCTTAACTTGGGAACTTATAGGGAGTTTACT 651
Db 65 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuGlySerPheThr 84
QY 652 GTTTTAGGAGAGGACACTCGTTGACTTTTCGAGAACATACGAGCTTCTCAAAATGGGCA 711
Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104
QY 712 GCTCTAAGTAATAGCGCTCTGATGAGCTGTTTACTATTGAGGGTTTTAAGAATTTATCC 771
Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124
QY 772 TTTTCCAATTCGAATTCATTACTTGGCGTACTGCTGCTGCAACGACTAATAAGGTAGC 831
Db 125 PheSerAsnCyAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144
QY 832 CAGACTCCGACGACACATCTACACCGCTCTAATGGTACTATTATTCTAAACAGACTCTT 891
Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 164
QY 892 TTGTACTCAATTAATGAGAAGTTCTCATTTCTATAGTAATTTAGTCTCTGAGATGGGGA 951
Db 165 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 184
QY 952 GCTATAGAGCTTAAGAGCTTAACGGTTCAAGGAATTTAGCAAGCTTTGTGTCTCCAGAA 1011
Db 185 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCyValPheGlnGlu 204
QY 1012 AATACTGCTCAAGCTATGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCT 1071
||||| 205 AsnThrAlaGlnAlaAspGlyGlyAlaCyGlnValValThrSerPheSerAlaMetAla 224
QY 1072 AACGAGGCTCTATTGCTTTGTAGCGAATCTTGCCAGGAGTAAGAGGGGAGGATGCT 1131
Db 225 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 244
QY 1132 GCTGTTCCAGGATGGCAGCAGGAGTGTCACTCATCTTCAACAAGAAGATCCAGTAGTA 1191
Db 245 AlaValGlnAspGlyGlnGlnGlyValSerSerThrThrGluAspProValVal 264
QY 1192 AGTTTTTCCAGAAATACCTGCGGTAGATTTCATGGGAACCTAGCCCGAGTAGGAGGG 1251
Db 265 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 284
QY 1252 ATTTTACTCTACGGGAGCTGTTCTCTGAATATGAAAAACCTGTTTCTCAACAAT 1311
Db 285 IleTyrSerTyrGlyAsnValAlaPheLeuAsnGlnGlyThrLeuPheLeuAsn 304
QY 1312 GTTGCTTCTCTGTTTACATTTGCTAAGCAACCAAGTGGACAGCTTCTTAATACG 1371
Db 305 ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 324
QY 1372 AGTAATAATTACGGAGATGGAGGACTATCTTCTGAAGATGCTGCGCAGCAGGATCC 1431
Db 325 SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 344
QY 1432 AATAACTCTGATCAGTTTCTTTCATCGAGAGGAGTAGTTTCTTTAGTAGCAATGTA 1491
Db 345 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 364
QY 1492 GCTGCTGGGAAAGGGGAGCTATTATATGCCAAAAAGCTCTCGGTTGCTAACTGTGCCCT 1551
Db 365 AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLeuSerValAlaAsnCyGlyPro 384
QY 1552 GTACAATTTTAAAGAAATATCGTAAATGATGTGTGAGCGATTATTATTAGGAGAACTGGA 1611
Db 385 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 404
QY 1612 GAGCTCAGTTTCTGCTGATTATCGAGATATTATTTTCGATGGGAATCTTAAAGAACA 1671
Db 405 GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 424
QY 1672 GCCAAGAGAAATGTCGCGATGTTAATGCGTAACCTGTGCTCAACAGCCATTTCCGATG 1731
Db 425 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 444
QY 1732 GGATCGGAGGGAATAAATACGACATTAAAGAGCTAAAGCAGGCGATCAGATTCTCTTAAT 1791
Db 445 GlySerGlyGlyLysIleThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 464
QY 1792 GATCCATCGAGATGGCAACCGAAATACACGAGCGAGCTCTTCCAACTTCTAAAA 1851
Db 465 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484
QY 1852 ATTAACGATGTGTAAGGATACACAGGGATATTGTTTTGCTTAATGGAAGCAGTACTTGTG 1911
Db 485 IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504
QY 1912 TACCAAAATGTTACGATAGACGAAGGATGTTCTTTCGTGAAAGCGAAATTTATCA 1971
Db 505 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 524
QY 1972 GTGAATTTCTTAAGTCAGACAGTGGAGTCTGTATATCGAAGCTGGAGTACATGGAT 2031
Db 525 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 544
QY 2032 TTTGTAACTCCCAACCAACCAACAGCTCTCTCGCGCTAAATCAGTTGATCAGCTTTTCC 2091
Db 545 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 564
QY 2092 AATCTGCATTTGCTCTTTCTTTCTTTGTTAGCAAAACAATGCAGTTACGAATCTCTTACC 2151
```

```
Db 565 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 584
Qy 2152 AATCCTCCAGCGCAAGATTCTCATCTCGAGTCATTTGGTAGCACAACTGCTGGTTCGTT 2211
Db 585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrAlaGlySerVal 604
Qy 2212 ACAATTAGTGGGCTATCTTTTGGAGATTGGATGATATACAGCTTATGATAGTAGAT 2271
Db 605 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaIleArgArgTyrAsp 624
Qy 2272 TGGCTAGGTTCTAATCAAAAATCAATGCTCTGAAATACAGTTAGGAGCTTAAGCCCCCA 2331
Db 625 TrpLeuGlySerAsnGlnIleAsnValIleuLysLeuGlnLeuGlyThrLysProPro 644
Qy 2332 GCTAATCCCCCATCAGATTGACTTAGGGAATCAGATGCTTAGTATGGCTATCAAGGA 2391
Db 645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly 664
Qy 2392 AGCTGGAAGCTTGGTGGGATCCTTAATACAGCAATTAATGGTCTTTATCTCTGAAGCT 2451
Db 665 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684
Qy 2452 ACATGGACTAAACTGGGTAAATCCTGGGCTCAGCGAGTAGCTCTTTGGTTCCAAAT 2511
Db 685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704
Qy 2512 AGTTATGGGATCCATTATAGATATACGATCTCGGCATTCAGCAATTCAGCAAGTGTG 2571
Db 705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
Qy 2572 GATGGGCGCTTATATGTCGAGGATATATGGTTCTGGAGTTTCGAATTTCTTATCAT 2631
Db 725 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPheTyrHis 744
Qy 2632 GACCGGATGCTTTAGTCAGGATATCGGTATATAGTGGGGTTATTCCTTAGGAGCA 2691
Db 745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
Qy 2692 AACTCCTACTTTGGATCATCATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751
Db 765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
Qy 2752 AAAGATTATGATGTGTGTTCCAAATCATCATGCTTCGATAGGATCCGTTTATCTATCT 2811
Db 785 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 804
Qy 2812 ACCCAACAGCTTTATGTCATCTTATTTGTCGAGATGCGTTTATCCGTGCTAGCTAC 2871
Db 805 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824
Qy 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTGCAGAGGAGCGATGTTGCT 2931
Db 825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844
Qy 2932 TGGGATAATAACTGTCTGGCTGGAGATTGGAGCGGATACCAATTGTGATTACTCCA 2991
Db 845 TrpAspAsnAsnCysLeuAlaGlyIleGlyAlaGlyLeuProIleValIleThrPro 864
Qy 2992 TCTAAGCTCTATTGAAAGATGTCCTTTTCGTCGAGCTGAGCTTCTTATCCCGAT 3051
Db 865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884
Qy 3052 CATGAATCTTTTACAGAGGAAGCGATCAAGCTCGGCGCATTCAGAGCGGACATCTCTTA 3111
Db 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
Qy 3112 AATCTATCAGTTCTCTGTTGGAGTGAAGTTTGATCGATGCTTAGTACATCCTAATAA 3171
Db 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924
Qy 3172 TATAGCTTTATCGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGCA 3231
Db 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
```

```
Qy 3232 ACCTCTATCCCATCAAGAGACATCGACAACAGATGCTTTTCATTTTAGCAAGACATGGA 3291
Db 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964
Qy 3292 GTTGTGTTAGAGATCTATGTATGCTTCTTCAACAAGTAATATAGAAGTATATGGCCAT 3351
Db 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984
Qy 3352 GGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTTGAGTGCAGGAAGTAGAGTC 3411
Db 985 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 1004
Qy 3412 CGGTTC 3417
Db 1005 ArgPhe 1006

RESULT 9
US-09-620-412C-190
; Sequence 190, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PR1
; ORGANISM: Chlamydia
US-09-620-412C-190

Alignment Scores:
Pred. No.: 0 Length: 1006
Score: 5090.00 Matches: 979
Percent Similarity: 99.9% Conservative: 2
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 64.6% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-620-412C-190 (1-1006)
Qy 472 ATGGTTCTCTCAAGCAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATATCTTT 531
Db 25 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 44
Qy 532 ATAGAGATCCGAGTGGGACTACTGTTTCTGCGAGAGATTAACTTAAAAATCTT 591
Db 45 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 64
Qy 592 GACAATTTCTATTCAGCTTTCCTTAAAGTTGTTTGGAACTTATTAGGAGTTTACT 651
Db 65 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 84
Qy 652 GTTTTAGGAGGAGGACACTCGTTGACTTTTCAGAAACATACCGACTTCTACAAATGGGCA 711
Db 85 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 104
Qy 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTATGAGGGTTTAAAGAATTATCC 771
Db 105 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 124
Qy 772 TTTTCCAAATTCGAATTCATTACTTCCCTGCTGCTGCAACCGACTTAATAGGCTAGC 831
Db 125 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 144
Qy 832 CAGACTCCGACGACCAACATCTACACCGTCTAATGCTACTATTATTCTTAAAAACAGATCTT 891
Db 145 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 164
```


QY	892	TTGTTACTCAATAATGAGAGTTCTCAATCTATAGTAATTTAGTCTCTGAGATCGGGA	951	QY	1972	GTGAATTTCTTAAGTCAGACAGGTGGGAGTCTGTATATGGAAGCTGGGAGTATCATGGGAT	2031
Db	165	LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly	184	Db	525	ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp	544
QY	952	GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTCTTCCAGAA	1011	QY	2032	TTTGTAATCTCCAAACACCACCAAGCTCTCGCGCTTAATCAGTTGATCAGCTTTCC	2091
Db	185	AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu	204	Db	545	PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer	564
QY	1012	AATACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCACCGTTTCTCTGCTATGGCT	1071	QY	2092	AATCTGCATTTGCTCTTCTTCTTTGTTAGCAACAATGCAGTTTACGAATCTCTCTACC	2151
Db	205	AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValThrSerPheSerAlaMetAla	224	Db	565	AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProThr	584
QY	1072	AACGAGCTCTATTCCCTTTGTAGCGAATGTGCAGGAGTAGAGGGGAGGATGGCT	1131	QY	2152	AATCTCCAGCCGCAAGATTCTCATCTGCAGTCATTGTGAGCACAACCTCTGTTCTGTT	2211
Db	225	AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla	244	Db	585	AsnProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal	604
QY	1132	GCTGTTTCAGATGGGAGGAGGTGTCATCATCTTCAACAGAAGATCCAGTAGTA	1191	QY	2212	ACATTAGTGGGCTTATCTTTTTCAGGATTTGGATGTAGTACAGCTTATCATAGTAGTATGAT	2271
Db	245	AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal	264	Db	605	ThrIleSerGlyProIlePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp	624
QY	1192	AGTTTTTCCAGAAATCTCGGTAGAGTTGTGGAACGTAGCCGAGTAGGAGGGG	1251	QY	2272	TGCTAGGTTCCTAATCAAAAAATCAATGTCTCTGAAATTAACAGTTAGGGACTAAGCCCCA	2331
Db	265	SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly	284	Db	625	TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro	644
QY	1252	ATTACTCTCAGCGGAACGTTGCTTCTCCTCAATAATGGAAAAACCTTGTCTCACAAT	1311	QY	2332	GCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCCCTAAGTATGCTATCAAGGA	2391
Db	285	IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn	304	Db	645	AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly	664
QY	1312	GTGCTTCTCCTGTTTACATGCTGTAGCAACCAACAGTGCACAGGCTCTTAATACG	1371	QY	2392	AGCTGAAAGCTGCGGGGATCCTTAATACAGCAATAATAGTCTTATCTCTGAAAGCT	2451
Db	305	ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr	324	Db	665	SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla	684
QY	1372	AGTAATAATPACGAGATGAGGAGCTATCTCTGTAAGAATGCTGCGCAAGCAGATCC	1431	QY	2452	ACATGCACTAAACCTCGGTATATCTCGGCCCTGACGAGTAGTCTTCTTGGTTCCAAAT	2511
Db	325	SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer	344	Db	685	ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn	704
QY	1432	ATAACTCTGGATCAGTTTCTTGTAGGAGGAGTAGTTTCTTCTTAGTCAATGTA	1491	QY	2512	AGTTTATGGGATCCCAATTTTAGATATACGATCTGCGCATTCAGCAATTCAGCAAGTGTG	2571
Db	345	AsnAsnSerGlySerValSerPheAspGlyGlyValValPhePheSerSerAsnVal	364	Db	705	SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal	724
QY	1492	GCTGCTGGGAAAGGGGAGCTATTATGCCAAAAAGCTCTCGGTGCTAACTGTGGCCCT	1551	QY	2572	GATGGGCTCTTATTGTGAGGATTTAGGTTTCTGGATTTCGAGTTTCGAATTTCTCTATCAT	2631
Db	365	AlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro	384	Db	725	AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis	744
QY	1552	GTACAAATTTTAAGGAATATCGCTAATGATGGTGGACCGATTATTATTAGGAAATCTGGA	1611	QY	2632	GACCGGATCTTTTAGTCTCAGGGATATCGGTATATTAGTGGGGGTATTCTCTTAGGAGCA	2691
Db	385	ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly	404	Db	745	AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyTyrSerLeuGlyAla	764
QY	1612	GAGCTCAGTTTATCTGCTGATTATGAGATATTATTTTCGATGGGAATCTTAAAGAAC	1671	QY	2692	AACTCTCTCTTTCGATCATCGATGTTTGGTCTAGCATTTTACCGAAGTATTGTTAGTATCT	2751
Db	405	GluLeuSerLeuSerAlaAspTyrGlyAspIlePheAspGlyAsnLeuLysArgThr	424	Db	765	AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer	784
QY	1672	GCCAAAGAGATCTCGGATGTTAATGGCGTAACTGTCTCCTCACAGGCATTTCCGATG	1731	QY	2752	AAAGATTATGATGTGCTGTTCCAAATCATCATCTGCTTGCATAGGATCCGTTTATCTATCT	2811
Db	425	AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet	444	Db	785	LysAspTyrValValCysArgSerAsnHisHisAlaCysIleGlySerValTyrLeuSer	804
QY	1732	GGATCGGAGGGAATAACACATTAAGAGCTAAAGCAGGCGCATCAGATTCTCTTTAAT	1791	QY	2812	ACCCAACAGCTTTATGTGATCTTATTTCTTGGAGATCGGTTCCTCGTCTAGCTAC	2871
Db	445	GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn	464	Db	805	ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr	824
QY	1792	GATCCCCATCGAGATGGCAACCGAAATAACACAGCCAGCGCATCTTCCAAACTTCTAAAA	1851	QY	2872	GGGTTTGGGAATCAGCATATGAAAACTCATATACATTTTCAGAGGAGAGCGATGTCGT	2931
Db	465	AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys	484	Db	825	GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg	844
QY	1852	ATTAACGATGCTGAGGATACACAGGGGATTTCTTTTGTCTAATGGAAGCAGTACTTG	1911	QY	2932	TGGGATTAATAACTGTCTGCTGGAGAGATTGGAGCGGATTTACCGATTGTGATTACTCCA	2991
Db	485	IleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerThrLeu	504	Db	845	TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro	864
QY	1912	TACCAAAATGTTACGATAGACAAAGGAGGATGTTCTTCTGTAAGGCGCAAAATATCA	1971	QY	2992	TCTAAGCTCTATTGATGATGTCGCTCTTCTCGTCAAGCTGAGTTTCTTATGCGCAT	3051
Db	505	TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer	524	Db	865	SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp	884
				QY	3052	CATGAATCTTTTACAGAGGAAGGCGATCAAGCTCGGCGATTTCAGAGCGGACATCTCTCTA	3111

465 AspProIleGluWeAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 484
 1852 ATTAACGATGTTGAAGGATACACAGGGGATATGTTTTCGCTAAATGGAAGCAGTACTTGG 1911
 485 IleAsnAepGlyGluGlyThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 504
 1912 TACCAAAATGTTACGATAGACAGCAAGCAAGATGTTCTTCGTGAAAGCCAAATATCA 1971
 505 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGlnLysAlaLysLeuSer 524
 1972 GTGAATTTCTTAAGTCAGACAGTGGGAGTCTGTATATGGAAGCTGGGAGTACATGGGAT 2031
 525 ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 544
 2032 TTTGTAACTCCAAACACCAACACAGCTCTCGCGCTAAATCAGTGTGATCAGCTTTCC 2091
 545 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 564
 2092 AATCTGCATTTGCTCTTCTTCTTCTTGTAGCAACAATGCAGTTACGAATCCTCTACC 2151
 565 AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAlaValThrAsnProProThr 584
 2152 AATCTCCAGCGCAAGATCTCATCTCGAGTCATTTGGTAGCACAACCTGCTGCTTCTGTT 2211
 585 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 604
 2212 ACAATTAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTATGATAGTATGAT 2271
 605 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 624
 2272 TGCTAGGTTCTTAATCAAAATCAATGCTCTGAAATACAGTTAGGACTAAGCCCCCA 2331
 625 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 644
 2332 GCTAATGCCCATCAGATTTGACTTAGGGAATGAGATGCTTAAGTATGGCTATCAAGGA 2391
 645 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 664
 2392 AGCTGGAAGCTTCGCTGGGATCTTAATACAGCAAAATAGTCTTATATCTCGAAGCT 2451
 665 SerTrpLysLeuAlaTyrAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 684
 2452 ACATGGAATAAACTGGGTAAATCTCGGCGCTGAGCGAGTACGCTTCTTGGTTCCAAAT 2511
 685 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 704
 2512 AGTTATGGGATCCATTTTAGATATACGATCGGCATTCGCAATTCAGCAAGGTG 2571
 705 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 724
 2572 GATGGGCGCTCTTATTCGAGGATATGAGTCTGAGTTCGAGTTTCGAAATTTCTTATCAT 2631
 725 AspGlyArgSerTyrCysArgLeuTyrValSerGlyValSerAsnPhePheTyrHis 744
 2632 GACCGCATCTTTAGTCAGGATATCGGTATATTAGTGGGTTATTCCTTAGGAGCA 2691
 745 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 764
 2692 AACTCTTACTTTGATCATCGATGTTGTTGCTAGCATTTACCGAAGTATTGGTAGATCT 2751
 765 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 784
 2752 AAAGATATTAGTGTGCTCGTTCCAAATCATCATCTGTCATGATCGGTTTATCTATCT 2811
 785 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 804
 2812 ACCCAACAGCTTTTATGAGATCTTATTTGTCGAGATGCGTTTATCCGCTAGCTAC 2871
 805 ThrGlnGlnAlaLeuCysArgLysTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 824
 2872 GGGTTTGGGATCAGCATATCAAACTCATATACATTTGCAAGGAGGAGGATGTCGT 2931

825 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 844
 2932 TGGGATATAACTGTCGCTGGAGAGATTCGAGCGGATACCGATTTGATGATTACTCCA 2991
 845 TrpAspAsnAsnCysLeuAlaGlyIleGlyAlaGlyLeuProIleValIleThrPro 864
 2992 TCTAAGCTCTATTTGAATGAGTTGCTCTTTCGTGCAAGCTGAGTTTCTTATCCCGAT 3051
 865 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 884
 3052 CATGAATCTTTTACAGAGAGGCGCATCAAGCTCGGCATTCAGAGCGGACATCTCCTA 3111
 885 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 904
 3112 AATCTATCAGTTCTCTGTTGGAGTGAAGTTTTCATGATGTTCTAGTACACATCTTAATAA 3171
 905 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 924
 3172 TATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGGTACTGAGACA 3231
 925 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 944
 3232 AGCTCTCTATCCCATCAAGACATGGACAACAGATGCCCTTTCATTTAGCAAGACATGGA 3291
 945 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 964
 3292 GTTGTGTTAGAGGATCTATGATGCTCTCTAAACAAGTAATATAGAAGTATATGGCCAT 3351
 965 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 984
 3352 CGAAGATATGATGATCGAGATGCTTCTCGAGGCTATGTTTTCAGTGCAGGAAGTAGATC 3411
 985 GlyArgTyrGluTyrArgAspAlaSerArgLysTyrGlyLeuSerAlaGlySerLysVal 1004
 3412 CGGTTC 3417
 1005 ArgPhe 1006

RESULT 11

US-09-556-877-176
 ; Sequence 176, Application US/09556877
 ; Patent No. 6432916
 ; GENERAL INFORMATION:
 ; APPLICANT: Probat, Peter
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Fling, Steve
 ; APPLICANT: Maisonneuve, Jeff
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C5
 ; CURRENT APPLICATION NUMBER: US/09/556,877
 ; CURRENT FILING DATE: 2000-04-19
 ; NUMBER OF SEQ ID NOS: 305
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 176
 ; LENGTH: 982
 ; TYPE: PRT
 ; ORGANISM: Chlamydia
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1) ... (982)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 ; US-09-556-877-176

Alignment Scores:
 Pred. No.: 0 Length: 982
 Score: 5084.00 Matches: 978
 Percent Similarity: 99.8% Conservative: 2
 Best Local Similarity: 99.6% Mismatches: 2
 Query Match: 64.5% Indels: 0
 DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-556-877-176 (1-982)

QY 472 ATGGTTCCTCAAGGAATTACGATGGGAGACGTTAACTGTATCAATTTCCCTATCTGTT 531
Db : : : : :
QY 1 MetIleProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyrThrVal 20
QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTTCGAGGAGAGTTAACTTAAATAATCTT 591
Db : : : : :
QY 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40
QY 592 GACAAATCTATTGCACTTTCCTTAAAGTTGTTTGGAACTTATTAGGAGATTTTACT 651
Db : : : : :
QY 41 AspAsnSerIleAlaAlaLeuProLysSerCysPheGlyAsnLeuLeuGlySerPheThr 60
QY 652 GTTTTGGGAGAGGACACTCGTTGACTTTCGAGAAACATACGAGCTTCTACAAATGGGCA 711
Db ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
QY 712 GCTCTAAGTAATAGCGCTGCTGATCGACTGTTTACTATTAGGGTTTAAAGAAATATCC 771
Db : : : : :
QY 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSer 100
QY 772 TTTTCCAAATGCAATTCATTACTCGCTGCTGCTGCTGCAACGACTAAATAAGGTAGC 831
Db : : : : :
QY 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120
QY 832 CAGACTCCGACGACAACTCTACACCGCTCTAAATGGTACTATTATTCTTAAACAGATCTT 891
Db GlnThrProThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 140
QY 892 TTGTTACTCAATAATAGAGAGTTCTCATCTATAGTAAATTAGTCTCTGGAGATCGGGA 951
Db : : : : :
QY 141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160
QY 952 GCTATAGATGCTAAGAGCTTCAAGGTTCAAGAAATTTAGCAAGCTTTGTCTTCCAAGAA 1011
Db AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180
QY 1012 AATACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCT 1071
Db AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200
QY 1072 AACGAGGCTCTATTGCTTTCCTAGCAATGTTGACAGGATGAGAGGGAGGATGCT 1131
Db : : : : :
QY 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 220
QY 1132 GCTGTTTCAGATGGGAGGAGGTCTCATCTACTTCAACAGAGATCCAGTAGTA 1191
Db AlaValGlnAspGlyGlnGlnGlyValSerSerSerSerThrGluAspProValVal 240
QY 1192 AGTTTTTCCAGAAATACTCGGCTAGAGTTTGTAGGAACTGACCGAGTAGGAGAGGG 1251
Db SerPheSerArgAsnThrAlaValAlaGluPheAspGlyAsnValAlaArgValGlyGly 260
QY 1252 ATTTACTCTACGGAACTGCTTTCCTGTAATATATGGAAAAACCTGTTTCTCAACAAT 1311
Db : : : : :
QY 261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 280
QY 1312 GTTGCTTCTCCTGTTTACATTCGCTAAGCAACCAAGTGGACAGGCTTCTAATACG 1371
Db ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300
QY 1372 AGTAATAATTACGGAGATGGAGAGCTATCTTCTGTAAGAATGGTCGCAAGCAGGATCC 1431
Db SerAsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320
QY 1432 AATAACTCTGGATCAGTTCTCTTGTATGAGAGGGAGTAGTTTCTTTAGTAGCAATGTA 1491
Db : : : : :
QY 321 AsnAsnSerGlySerValSerPheAspGlyGlyGlyValValPhePheSerSerAsnVal 340
QY 1492 GCTGCTGGGAAAGGGAGCTATTTATGCAAAAAGCTCTCGTTGCTAACTGTGGCCCT 1551
Db AlaAlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360

QY 1552 GTACAATTTTAAAGAAATATCGCTAATGATGGTGGAGGATTTATTATAGGAGATCTCGA 1611
Db ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGlySerGly 380
QY 1612 GAGCTCAGTTTATCTGCTGATTATGGAGATATATTTTCGATGGAAATCTTAAAGAACAA 1671
Db GlnLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400
QY 1672 GCCAAGAGAAATGCTGCCGATGTTAAATGGCTAACTGTGTCTCTCAACAGCCATTTCCGATG 1731
Db AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420
QY 1732 GGATCGGAGGAGAAATAACGACATTAAGACTAAACGAGGCGATCAGATTCTCTTTAAT 1791
Db GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440
QY 1792 GATCCATCGAGATGGCAACGAAATAACACGAGCCAGCGAGTCTTCCAAAATCTTCAAAA 1851
Db AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 460
QY 1852 ATTAACGATGTTGAAGGATACACAGGGGATATTGTTTTCCTTAATGGAAGCAGTAGTCTTG 1911
Db IleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
QY 1912 TACCAAAATGTCATAGACGAAAGAGATTGTTTCTTCTGTAAGAAAGCGCAAAATATCA 1971
Db TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500
QY 1972 GTGAATTCCTTAAGTCAGACAGGTGGGAGTCTGTATATGGAAGCTGGGAGTACATGGAT 2031
Db ValAsnSerLeuSerGlnThrGlyLysLeuTyrMetGluAlaGlySerThrLeuAsp 520
QY 2032 TTTGTAACTCCACAACACCAACAGCTCTCTCGCGCTAAATCACTGATGATCAACGCTTTC 2091
Db PheValThrProGlnProGlnProAlaAlaAsnGlnLeuIleThrLeuSer 540
QY 2092 AATCTGCAATTTGCTCTCTTCTTCTTGTAGCAAAACATGAGTGTACAAATCTCTCTACC 2151
Db AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560
QY 2152 AATCTCCAGCGCAGGATTTCTCATCTCGAGTCATTGTTAGCACAACTGCTGGTTCTGTT 2211
Db AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580
QY 2212 ACAATTAGTGGCCCTATCTTTTGGAGATTGGATGATACAGCTTATGATAGGTATGAT 2271
Db ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 600
QY 2272 TGGCTAGGTTCTAATCAAAAATCAATCTCTGAAATTTACAGTATAGGAGCTAAGCCCCA 2331
Db TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
QY 2332 GCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCCCTAAGTATGCTATCAAGA 2391
Db AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyGlnGly 640
QY 2392 AGCTCGAAGCTTTCGCTGGGATCCTAATACAGAAAATAATGTCCTTATATCTCGAAGCT 2451
Db SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660
QY 2452 ACATCGACTAAACCTGGGTATATCTCGGCCCTGACGAGTAGTCTCTTGGTTCCAAAT 2511
Db ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
QY 2512 AGTTTATGGGATCCATTTTAGATATACGATCGCGCATTCAGCAATTCACAGCAAGTGTG 2571
Db SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
QY 2572 GATGGGCGCTCTTATTGTCGAGGATTTATGGGTTTCTCGAGTTTCGAATTTCTTCTATCAT 2631
Db AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720

```
QY 2632 GACCGGATGCTTTAGTTCAGGAGATATCGGTATATAGTGGGGTTATTCTTCAGGACCA 2691
Db |||||||
QY 721 AspA-gAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740
Db |||||||
QY 2692 AACTCTCTCTTTGGATCATCGATGTTGGTCTAGCATTTACCGAAGTATTGGTAGATCT 2751
Db |||||||
QY 741 AsnSerTyr-PheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
Db |||||||
QY 2752 AAAGATTATGATGTGTCTCCATCATCATCTGCTGATAGGATCCGTTTATCTATCT 2811
Db |||||||
QY 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
Db |||||||
QY 2812 ACCCAACAAGCTTTATGTGATCCTATTGTTCCGAGATGCGTTTATCCGTGCTAGCTAC 2871
Db |||||||
QY 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800
Db |||||||
QY 2872 GGGTTTGGGAATCAGCATATGAACCTCATATACATTTTCAGAGAGAGCGGATGTCGT 2931
Db |||||||
QY 801 GlyPheGlyAsnGlnHisMetLysThrSerTyr-ThrPheAlaGluGluSerAspValArg 820
Db |||||||
QY 2932 TGGGATAATAAATCTGCTCGCTGGAGAGATTTGGACGGGATTACCGATTGTGATTACTCCA 2991
Db |||||||
QY 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
Db |||||||
QY 2992 TCTAAGCTCTATTGATGATGCTGCTCTTCGTCGCAAGCTGAGTTTCTTATGCGGAT 3051
Db |||||||
QY 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
Db |||||||
QY 3052 CATGAATCTTTTACAGAGGAAGCGGATCAAGCTCGGGCAATTCAGAGCGGACATCTCTTA 3111
Db |||||||
QY 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
Db |||||||
QY 3112 AATCTATCAGTCTCTGTTGGAGTGAAGTTTGATGCGATGTTCTAGTACACATCTTAATAA 3171
Db |||||||
QY 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
Db |||||||
QY 3172 TATAGCTTTATGCGGCTTATCTGATGCTTATGCGACCATCTCTGCTACTGAGACA 3231
Db |||||||
QY 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920
Db |||||||
QY 3232 ACCTCTCTATCCCATCAAGAGACATGGACACAGATGCCCTTCAATTTAGCAAGACATGGA 3291
Db |||||||
QY 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940
Db |||||||
QY 3292 GTTGTGGTATAGAGATCTATGATGCTTCTCTCAAGATATATAGAGTATATGCGCAT 3351
Db |||||||
QY 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960
Db |||||||
QY 3352 GGAAGATATAGTATCGAGATGCTTCTCGAGGCTATGTTGATGTCGAGGAGTAGATC 3411
Db |||||||
QY 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980
Db |||||||
QY 3412 CGGTTT 3417
Db |||
QY 981 ***Phe 982
```

RESULT 12

US-09-620-412C-176

; Sequence 176, Application US/09620412C

; Patent No. 6448234

; GENERAL INFORMATION:

; APPLICANT: Steven P. Fling

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C7

; CURRENT APPLICATION NUMBER: US/09/620.412C

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 176

; LENGTH: 982

; TYPE: PRT

; ORGANISM: Chlamydia

```
! FEATURE:
! NAME/KEY: VARIANT
! LOCATION: (1)---(982)
! OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-412C-176
```

```
Alignment Scores:
Pred. No.: 0 Length: 982
Score: 5084.00 Matches: 978
Percent Similarity: 99.8% Conservative: 2
Best Local Similarity: 99.6% Mismatches: 2
Query Match: 64.5% Indels: 0
DB: 2 Gaps: 0
```

US-10-701-844-1 (1-4435) x US-09-620-412C-176 (1-982)

```
QY 472 ATGGTTCCTCAAGGAATTTACGATGGGAGAGCTTAACCTGATCATCTTCCCTATACGT 531
Db |||||||
QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTCTCGCAGGAGAGTTAACATTAAAAATCTT 591
Db |||||||
QY 21 IleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40
Db |||||||
QY 592 GACAAATCTATTGCAGCTTTGCTTTTAAGTTGTTTGGCACTTATTAGGGAGTTTACT 651
Db |||||||
QY 41 AspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuLeuGlySerPheThr 60
Db |||||||
QY 652 GTTTTAGGAGAGGACACTCGTTGACTTTTCGAGACATACGGACTTCTCAATGGGCA 711
Db |||||||
QY 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
Db |||||||
QY 712 GCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTATTGAGGTTTAAAGAAATATCC 771
Db |||||||
QY 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGlyPheLysGluLeuSer 100
Db |||||||
QY 772 TTTTCCAAATGCAATTCATTACTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
Db |||||||
QY 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120
Db |||||||
QY 832 CAGACTCCGAGCAGACATCTACCGCTCAATGCTACTATTATTTCTTAAACAGATCTT 891
Db |||||||
QY 121 GlnThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeu 140
Db |||||||
QY 892 TTGTTACTCAATAATCAGAGTTCTCATCTCTAGTAATTTAGTCTCTGAGAGATGGGGA 951
Db |||||||
QY 141 LeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGly 160
Db |||||||
QY 952 GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTTAGCAAGCTTTGTGCTTCCAAGAA 1011
Db |||||||
QY 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180
Db |||||||
QY 1012 AATPACTGCTCAAGCTGATGGGAGCTTGTCAAGTAGTCAACAGTTTCTCTGCTATGGCT 1071
Db |||||||
QY 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200
Db |||||||
QY 1072 AACGAGCTCTTATTCCTTTGTAGCGAATGTTGCGAGGCTTAAGAGGGGAGGATGCT 1131
Db |||||||
QY 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyIleAla 220
Db |||||||
QY 1132 GCTGTTTCAGGATGGGAGGAGGAGTGTCTATCTACTTCAACAGAAAGATCCAGTAGTA 1191
Db |||||||
QY 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrSerThrGluAspProValVal 240
Db |||||||
QY 1192 AGTTTTTCCAGAAATACTCGCGTAGATGTTGATGGGAACCTAGCCCGAGTAGGAGGAGG 1251
Db |||||||
QY 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGly 260
Db |||||||
QY 1252 ATTTACTCTTACGGGAACGTTGCTTCTCAATTAATCGAAAAACCTTGTCTTCAACAAT 1311
Db |||||||
QY 261 IleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsnAsn 280
Db |||||||
QY 1312 GTTGCTTCTCTCTGTTTACATTGCTGCTAAGCAACCAAGTGCAGAGGCTTCTTAATACG 1371
Db |||||||
```

Db ValAlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300
QY AGTAATAATTACGGAGATGGAGGACTATCTTCTGAAGAATGGTGGCCAGCAGGATCC 1431
Db SerAsnAsnTyrGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320
QY AATAACTCTGATCAGTTCTCTTTCATCGAGAGGAGTAGTTCTTCTTAGTACCAATGTA 1491
Db AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 340
QY GCTGCTGGAAAGGGGAGCTATTATGCCAAAAGCTCTGGTGTAACTGGCCCT 1551
Db AlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyPro 360
QY GTACAAATTTTAAAGGAATATCGCTAATGATGTGGAGCGATTATTTAGGAGAATCTGGA 1611
Db ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyrLeuGlyGluSerGly 380
QY GAGCTCAGTTTATCTGCTGATATCGAGATATTTATTTTCGATGGGAATCTTAAAGAACA 1671
Db GluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400
QY GCCAAGAGATGCTGCCGATGTTAATCGCGTAACGTGTCTCAAGCCATTTCCGATG 1731
Db AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420
QY GGATCGGAGGAGAAATAACGACATTAAGAGCTAAAGCAGGGCATCAGATCTCTTTAAT 1791
Db GlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440
QY GATCCATCGAGATGGCAACGGAATAAACCGCCAGCGCAGTCTTCCAAACTTCTAAAA 1851
Db AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLys 460
QY ATTAACGATGTGAAGGATACACAGCGATATTTGTTTGTAAATGGAAGCGATCTTGG 1911
Db IleAsnAspGlyGluGlyTyrThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
QY TACCAAAATGTTACCATAGACAGCAAGGATGTTCTTGTGAAAGGCAAAATTTATCA 1971
Db TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500
QY GTGAATTTCTTAAGTCAGACAGTGGAGTCTGTATATGGAAGCTGGGAGTACATGGAT 2031
Db ValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGlySerThrLeuAsp 520
QY TTTGTAACTCCACCAACCAACCAAGCTCTCGCGCTATCATCAGTTGATCAGCTTTCC 2091
Db PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIleThrLeuSer 540
QY AATCTGCAATTTCTCTTCTTCTTCTTGTGTAGCAAAATGCAAGTTACGAATCTCTACC 2151
Db AsnLeuHisLeuSerLeuSerSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560
QY AATCTCCAGCCCAAGATTTCTATCTCAGTCATTGGTGGTGGTGGTGGTGGTGGTGGT 2211
Db AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580
QY ACAATAGTGGCCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGTATGAT 2271
Db ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyrAspArgTyrAsp 600
QY TGGCTAGGTTCTTAATCAAAAATCAATGTCCTGAAATACAGTTAGGGAAGTAAAGCCCA 2331
Db TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
QY GCTAATGCCCCATCAGATTGATCTCTAGGGAATGAGATGCCCTAAGTATGGCTATCAAGA 2391
Db AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGly 640
QY AGCTGGAGCTTGGTGGATCTTAATACAGCAAAATAGTGTCTTATCTCTGAAGCT 2451

Db SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660
QY ACATGGACTAAACTGGGTATAATCTGGGCTCAGCGAGTAGCTTCTTTGGTTCCAAAT 2511
Db ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
QY AGTTATGGGATCCATTTTAGATATACGATCTGGCATTCAGCAATTCAGCAAGTGTG 2571
Db SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
QY GATGGGCTCTTATTTGTGAGGATATGGGTTCCTGGAGTTTCGAAATTTCTTCTATCAT 2631
Db AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720
QY GACCGCATGCTTTAGGTCAGGGATATCGGTATATTAGTGGGGTTATCTCTTAGGACA 2691
Db AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740
QY AACTCTACTTTGATCATCGATGTTGGTCTAGCATTTACCGAGTATTTGGTAGATCT 2751
Db AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
QY AAGATATGATGATGTGCTGTTCCAAATCATCATGCTTCATAGGATCCGTTTATCTATCT 2811
Db LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
QY ACCCAACAGCTTTTATGTTGATCTCTTATTTGTTGGAGATGCGTTTATCCGTGCTAGTAC 2871
Db ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800
QY GGGTTTGGGATCAGCATATGAAAACCTCATATATACATTTGCAGAGGAGCGATTTGCT 2931
Db GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820
QY TGGGATAATACTGTCTGGCTGGAGAGATTCGAGCGGATTCACGATTTGATTAATCTCA 2991
Db TrpAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
QY TCTAAGCTCTATTGAAATGAGTTGCTTCTTCTGCGAAGCTGAGTGTCTTATATGCCAT 3051
Db SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
QY CATGAATCTTTTACAGAGAGGCGATCAAGCTCGGCGATTCAGAGCGGACATCTCCTA 3111
Db HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
QY AATCTATCAGTTCTGTTGGAGTGAAGTTTCATCGATGTTCTAGTACATCTCTAATAAA 3171
Db AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
QY TATAGCTTTATGGCGCTTATATCTGTGATGCTTATCGCACATCTCTGTACTGAGACA 3231
Db TyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSerGlyThrGluThr 920
QY AGCTCTCTATCCCATCAAGACATGAGACACAGATGCTTTCATTTAGCAAGACATGGA 3291
Db ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940
QY GTTGTGGTTAGAGGATCTATGTTATGCTTCTTAAACAAGTAATATAGAGTATATGGCCAT 3351
Db ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960
QY GGAAGATGATGATTCGAGATGCTTCTCGAGCTATGTTTGTAGTGCAGGAAGTAGAGTC 3411
Db GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980
QY CGGTTTC 3417
Db ***Phe 982

RESULT 13
US-09-598-419-176
; Sequence 176, Application US/09598419

Patent No. 6565856
 GENERAL INFORMATION:
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Scholler, John
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
 TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
 FILE REFERENCE: 210121.469C6
 CURRENT APPLICATION NUMBER: US/09/598,419
 CURRENT FILING DATE: 2000-06-20
 NUMBER OF SEQ ID NOS: 357
 SOFTWARE: FastSeq for Windows Version 3.0/4.0
 SEQ ID NO 176
 LENGTH: 982
 TYPE: PRT
 ORGANISM: Chlamydia
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)...(982)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-598-419-176

Alignment Scores:
 Pred. No.: 0 Length: 982
 Score: 5084.00 Matches: 978
 Percent Similarity: 99.8% Conservative: 2
 Best Local Similarity: 99.6% Mismatches: 2
 Query Match: 64.5% Indels: 0
 DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-598-419-176 (1-982)

QY 472 ATGGTTCCTCAAGGATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATACGTT 531
 DB 1 MetIleProGlnGlyIleTyAspGlyGluThrLeuThrValSerPheProTyThrVal 20
 QY 532 ATAGGAGATCCGAGTGGGACTACTGTTTCTCGCAGGAGAGTTAAACATTAATAATCTT 591
 DB 21 IleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLysAsnLeu 40
 QY 592 GACAAATCTATTCAGCTTTGCTTTTAAAGTTGTTGGGAACTTATAGGAGTTTACT 651
 DB 41 AspAsnSerIleAlaLeuProLeuSerCysPheGlyAsnLeuLysGlySerPheThr 60
 QY 652 GTTTTAGGAGAGACACCTGTTGACTTTCGAGAACATAGGACTTCTACAAATGGGCA 711
 DB 61 ValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThrAsnGlyAla 80
 QY 712 GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTTATTCAGGCTTTTAAAGAAATATCC 771
 DB 81 AlaLeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGlyPheLysGluLeuSer 100
 QY 772 TTTTCCAAATTCGAATTCATTACTTTCGCTACTGCTGCTGCAACGACTAATAAGGGTAGC 831
 DB 101 PheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySer 120
 QY 832 CAGACTCCGAGGACACATCTACACCGTCTAATGGTACTATTATTTCTAAACAGACTCT 891
 DB 121 GlnThrProThrThrSerThrProSerAsnGlyThrIleTySerLysThrAspLeu 140
 QY 892 TTGTTACTCAATATAGAGATGTTCTCAATCTATAGTAATTTAGTCTCTGGAGATGGGGA 951
 DB 141 LeuLeuLeuAsnAsnGluLysPheSerPheTySerAsnLeuValSerGlyAspGlyGly 160
 QY 952 GCTATAGATGCTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAA 1011
 DB 161 AlaIleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysValPheGlnGlu 180
 QY 1012 AATCTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTACAGTCTTCTGCTATGCT 1071
 DB 181 AsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAla 200
 QY 1072 AACGAGGCTCTATGCTTTGTAGCGAATGTTGCAGGAGTTAAGAGGGGAGGATGCT 1131

DB 201 AsnGluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAla 220
 QY 1132 GCTGTTTCCAGGATGGCAGCAGGAGTGTCACTCATCTTCAACAGAGATCCAGTAGTA 1191
 DB 221 AlaValGlnAspGlyGlnGlnGlyValSerSerSerThrThrGluAspProValVal 240
 QY 1192 AGTTTTTCCAGAAATPACTCGCGTAGCTTTGATGGGAACGTAGCCGAGTAGGAGAGG 1251
 DB 241 SerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyGly 260
 QY 1252 ATTTACTCTACGGGAAAGTTGCTTCTCTGATATATGGAACAACTTCTTCTCAACAT 1311
 DB 261 IleTySerTySerGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPheLeuAsn 280
 QY 1312 GTTGCTTCTCTGTTTACATTCCTAAGCAACCAACAAAGTGGCAGCAGCTTCTAATACG 1371
 DB 281 ValAlaSerProValTyIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThr 300
 QY 1372 ACTAATAATTACGGAGATGGAGGACTATCTTCTGTAAGAATGGTGGCAGCAGGATCC 1431
 DB 301 SerAsnAsnTyGlyAspGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySer 320
 QY 1432 AATAACTCTGGATCATGTTTCTTCTGATGGAGAGGAGTAGTTTCTTTAGTACCAATGA 1491
 DB 321 AsnAsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnVal 340
 QY 1492 GCTGCTGGGAAAGGGGAGCTATTATGCAAAAAGCTCTCGGTGCTAATCTGTGCCCCT 1551
 DB 341 AlaAlaGlyLysGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCysGlyPro 360
 QY 1552 GTACAAATTTTAAAGGAATATCCTAATGATGTGGAGCGCATTTATTTAGAGAAATCTGGA 1611
 DB 361 ValGlnPheLeuArgAsnIleAlaAsnAspGlyGlyAlaIleTyLeuGlyGluSerGly 380
 QY 1612 GAGCTCAGTTTATCTGCTGATTAAGGATATATTTTCGATGGGATCTTAAAGAACACA 1671
 DB 381 GluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsnLeuLysArgThr 400
 QY 1672 GCCAAGAGAACTGTCGCGATGTTAATGCGTAACGTCTCTCACAAGCCATTTCCGATG 1731
 DB 401 AlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMet 420
 QY 1732 GGATCCGGGAGGAAAAATAACGACATTAAGAGCTAAAGCAGGCGCATCAGATTCCTTAAT 1791
 DB 421 GlySerGlyGlyLysIleThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsn 440
 QY 1792 GATCCCATCAGATGGCAACCGAAATAACCGACCGACCGCAGCTCTTCCAACTCTAAA 1851
 DB 441 AspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerLysLeuLeuLys 460
 QY 1852 ATTAACGATGGTGAAGGATACACAGGGGATATTGTTTTGCTAATGGAAGCAGTACTTTG 1911
 DB 461 IleAsnAspGlyGluGlyTyThrGlyAspIleValPheAlaAsnGlySerSerThrLeu 480
 QY 1912 TACCAAAATGTTACGATAGACGAGGAGGATTTCTTCTCGTGAAGGCAAAATATATCA 1971
 DB 481 TyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSer 500
 QY 1972 GTCAATTCCTAAGTCAGACAGCTGGAGTCTGTATATGGAAGCTGGAGTACATGGAT 2031
 DB 501 ValAsnSerLeuSerGlnThrGlySerLeuTyMetGluAlaGlySerThrLeuAsp 520
 QY 2032 TTTGTAACCTCCCAACCAACCAACAGCTCTCGCGCTAATCATGTTGATCAGCTTTCC 2091
 DB 521 PheValThrProGlnProGlnGlnProProAlaAlaAsnGlnIleThrLeuSer 540
 QY 2092 AATCTGCAATTTGCTCTTCTTCTTCTTCTAGCAAAACAAATCAGTTACGAATCTCTTACC 2151
 DB 541 AsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThr 560
 QY 2152 AATCTCTCAGCGCAAGATCTCATCTGCGAGTCAATGCTGAGCAACATGCTGTTGTT 2211
 DB 561 AsnProProAlaGlnAspSerHisProAlaValIleGlySerThrThrAlaGlySerVal 580

QY 2212 ACAATTAGTGGCCCTATCTTTTTCAGGATTTGGATGATACAGCTTATGATAGGTATGAT 2271
 Db 581 ThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyAspArgTyrAsp 600
 QY 2272 TGGCTAGGTTCTTAATCAAAAATCAATCTCCTGAAATTTACAGTTAGGAGCTAAAGCCCA 2331
 Db 601 TrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeuGlyThrLysProPro 620
 QY 2332 GCTAATGCCCATCAGATTGACTCTAGGGAATGAGATGCCCTAAGTATGGCTATCAAGGA 2391
 Db 621 AlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyThrGlnGly 640
 QY 2392 AGCTGGAAGCTTGGCTGGATCTTAATACAGCAAAATATGGTCTTATCTCTGAAAGCT 2451
 Db 641 SerTrpLysLeuAlaTrpAspProAsnThrAlaAsnAsnGlyProTyrThrLeuLysAla 660
 QY 2452 ACATGGACTAAACTGGGTATATCTGGCTGAGCGAGTAGCTTCTTGGTCCCAAT 2511
 Db 661 ThrTrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsn 680
 QY 2512 AGTTTATGGGATCCATTTAGATATACGATCTGGCATTACGCAATTCAGCAAGTGTG 2571
 Db 681 SerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerVal 700
 QY 2572 GATGGCGCTCTTATTTGTCGAGGATATGGGTTCTGGAGTTTCGAATTTCTTCTATCAT 2631
 Db 701 AspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPhePheTyrHis 720
 QY 2632 GACCGCGATCTTTAGTTCAGGATATCGGTATATAGTGGGGTTATTCCTTAGGACA 2691
 Db 721 AspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAla 740
 QY 2692 AACTCTCTTCTGGATCATCGATGTTTGGTCTAGCATTTACCGAAGTATTTGGTAGATCT 2751
 Db 741 AsnSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 760
 QY 2752 AAAGATTATGATGTGTGTTCCCAATCATCATCATGCTTGCATAGGATCCGTTTATCTACT 2811
 Db 761 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyrLeuSer 780
 QY 2812 ACCCAACAAGCTTTATGTGATCCTATTTGTTGCGGAGATGCGTTTATCGGTAGCTAC 2871
 Db 781 ThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAspAlaPheIleArgAlaSerTyr 800
 QY 2872 GGGTTTGGGAATCAGCATATGAAACCTCATATACATTTGCGAGAGAGCGCATGTCGT 2931
 Db 801 GlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGluSerAspValArg 820
 QY 2932 TGGGATAATACTGTCTGGCTGGAGATTTGAGCGGGATTACCGAATTTGATTACTCCA 2991
 Db 821 TrpAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIleValIleThrPro 840
 QY 2992 TCTAAGCTCTATTGATGATGTTGCTCTTTGCGCAAGCTGAGTTTCTTATCCGAT 3051
 Db 841 SerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAsp 860
 QY 3052 CATGAATCTTTTACAGAGAAAGCGCATCAAGCTCGGGCATTCAGAGCGGCATCTCCTTA 3111
 Db 861 HisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSerGlyHisLeuLeu 880
 QY 3112 AATCTATCAGTTCTCTGAGTGAAGTTTGATCGATGTTCTAGTACACATCCTAATAAA 3171
 Db 881 AsnLeuSerValProValGlyValLysPheAspArgCysSerSerThrHisProAsnLys 900
 QY 3172 TATAGTTTATCGCGCTTATATCTGTGATGCTTATCGCACCATCTCTGTTACTGAGCA 3231
 Db 901 TyrSerPheMetAlaAlaTyrIleCysAspAlaTyArgThrIleSerGlyThrGluThr 920
 QY 3232 ACGCTCTATCCATCAAGAGACATGCAACAGATGCTCTTCAATTTAGCAAGATGGA 3291
 Db 921 ThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeuAlaArgHisGly 940

QY 3292 GTTCTGCTAGAGCATCTATGCTATGCTCTCTAAAGTAATATAGAGTATATGCGCAT 3351
 Db 941 ValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGluValTyrGlyHis 960
 QY 3352 GGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3411
 Db 961 GlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAlaGlySerLysVal 980
 QY 3412 CGGTTTC 3417
 Db 981 ***Phe 982
 RESULT 14
 US-10-197-220-169
 ; Sequence 169, Application US/10197220
 ; Patent No. 6919187
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatic, Ajay
 ; APPLICANT: Guderian, Jeff
 ; APPLICANT: Skeiky, Yasir A. W.
 ; APPLICANT: Maisonneuve, Jean-Francois L.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515C3
 ; CURRENT APPLICATION NUMBER: US/10/197,220
 ; CURRENT FILING DATE: 2002-07-15
 ; NUMBER OF SEQ ID NOS: 175
 ; SEQ ID NO 169
 ; LENGTH: 670
 ; TYPE: PRT
 ; ORGANISM: Chlamydia
 US-10-197-220-169
 Alignment Scores:
 Pred. No.: 1,41e-300 Length: 670
 Score: 3336.50 Matches: 651
 Percent Similarity: 98.8% Conservative: 11
 Best Local Similarity: 97.2% Mismatches: 7
 Query Match: 42.3% Indels: 1
 DB: 2 Gaps: 1
 US-10-701-844-1 (1-4435) x US-10-197-220-169 (1-670)
 QY 463 GCAGAAATCATGTTCTCTCAAGGAATTTACATGGGAGAGCTTAACTGTATCATTTCCC 522
 Db 1 AlaGluIleMetIleProGlnGlyTyrAspGlyGluThrLeuThrValSerPhePro 20
 QY 523 TATACGTATTATAGGAGATCCGAGTGGGACTACTGTTTTCTTCGAGGAGAGTTAAACATTA 582
 Db 21 TyrThrValIleGlyAspProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeu 40
 QY 583 AAAAATCTTGACAATTTCTATTCAGCTTTGCTTTTAACTTTGTTTGGGAACCTTATTAGG 642
 Db 41 LysAsnLeuAspAsnSerIleAlaAlaLeuProLeuSerCysPheGlyAsnLeuGly 60
 QY 643 AGTTTATCTGTTTATAGGAGAGGACACTCGTTGACTTTTCGAGAAACATACGAGCTTCTACA 702
 Db 61 SerPheThrValLeuGlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr 80
 QY 703 AATGGGCGAGCTCTAAGTAATAGCGCTGCTGATGAGCTGTTTACTTATTGAGGGTTTAAA 762
 Db 81 AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGlyLeuPheThrIleGluGlyPheLys 100
 QY 763 GAATTTATCTTTTCCAAATTTGCAATTCATTTACTTTCGCTGCTGCTGCTGCTGCTGCTGCT 822
 Db 101 GluLeuSerPheSerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsn 120
 QY 823 AAGGTAGCCGAGCTCCGACGACCACTACACCGCTCTAATGGTACTTATTTATTTCTAAA 882
 Db 121 AsnGlySerGlnThrProThrThrThrSerThrProSerAsnGlyThrThrIleThrSerLys 140
 QY 883 ACAGATCTTTTCTTACTCAATAATCAGAAAGTTCTCTATTCTATAGTAATTTAGTCTCTGGA 942

Db 141 ThrAspLeuLeuLeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGly 160
Qy 943 GATGGGGAGCTATAGATCTTAAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTCTC 1002
Db 161 AspGlyGlyAlaAlaAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuCysVal 180
Qy 1003 TTCCCAAGAAATACTGCTCAAGCTCATGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCT 1062
Db 181 PheGlnGluAsnThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSer 200
Qy 1063 GCTATGGCTAACAGAGCTCTATTCGCTTTGTAGGAATGTTGCGAGGTAAGAGGGGA 1122
Db 201 AlaMetAlaAsnGluAlaProIleAlaPheIleAlaAsnValAlaGlyValArgGlyGly 220
Qy 1123 GGGATTGCTGTTTCAGATGGCGAGGAGGTGTCTATCTATCTTCAACAGAGAT 1182
Db 221 GlyIleAlaAlaValGlnAspGlyGlnGlnGlyValSerSerThrThrGluAsp 240
Qy 1183 CCAGTAGTAAGTTTTTCCAGAAATACTGCGGTAGAGTTTGTAGGGAACGTAGCCGAGTA 1242
Db 241 ProValValSerPheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgVal 260
Qy 1243 GGAGGAGGATTACTCTACGGGAACGTTGCTTTCTCTGAATAAGTAAAGAAACCTTGTCTT 1302
Db 261 GlyGlyGlyIleTyrSerTyrGlyAsnValAlaPheLeuAsnAsnGlyLysThrLeuPhe 280
Qy 1303 CTCACAAGTTGCTTCTCTGTTTACATGCTGTGCTAAGCAACCAACAGTGCAGAGCT 1362
Db 281 LeuAsnAsnValAlaSerProValTyrIleAlaAlaGluGlnProThrAsnGlyGlnAla 300
Qy 1363 TCTATACAGTAATAATACCGAGATGCGAGAGCTATCTCTGTAAAGATGTGCGCAA 1422
Db 301 SerAsnThrSerAspAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGln 320
Qy 1423 ---GCAGGATCCAAATACCTGATCAGTTTCTCTTGTAGGAGGAGTAGTTTCTTTT 1479
Db 321 AlaAlaGlySerAsnAsnSerGlySerValSerPheAspGlyGluGlyValPhePhe 340
Qy 1480 AGTAGCAATGTAGCTCTGGGAAAGGGGAGCTATTATTCGCAAAAAGCTCTCGTGTCT 1539
Db 341 SerSerAsnValAlaAlaGlyLysGlyAlaIleTyrAlaLysLysLeuSerValAla 360
Qy 1540 AACTGTGGCCCTGCAATTTTAAAGGAATATCGCTAATGATGCTGAGCGAGTATTATTTA 1599
Db 361 AsnCysGlyProValGlnPheLeuGlyAsnIleAlaAsnAspGlyGlyAlaIleTyrLeu 380
Qy 1600 GGAGAACTCGAGAGCTCAGTTTATCTGCTGATTATGAGATATTTTTCGATGGGAAT 1659
Db 381 GlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsn 400
Qy 1660 CTTAAAGAACACCCAAAGAAATGCTGCGATGTTAATGGCGTAATCTGTCTCTCACAA 1719
Db 401 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 420
Qy 1720 GCATTTTCGATGGATGGGAGGGAATAACACATTAAGACTAAGCAGGAGGATCAG 1779
Db 421 AlaIleSerMetGlySerGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln 440
Qy 1780 ATTCTCTTTAATCATCCATCGAGATGGCAACCGAAATAACCGAGCGCAGTCTTCC 1839
Db 441 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 460
Qy 1840 AAACTTTCTAAAAATTAACTGATGGTGAAGGATACACAGGGGATATGTTTGTCTAATGGA 1899
Db 461 GluProLeuLysIleAsnAspGlyGlyGlyTyrThrGlyAspIleValPheAlaAsnGly 480
Qy 1900 AGCAGTACTTTGACCAAAATGTTACATAGCAAGGAGGATGTTCTTCTCGTGAAG 1959
Db 481 AsnSerThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuAa-gluLys 500
Qy 1960 GCAAAATTATCAGTGAATTCCTTAAGTCAGACAGTGGGAGTCTGTATATGGAAGCTGGG 2019
Db 501 AlaLysLeuSerValAsnSerLeuSerGlnThrGlyGlySerLeuTyrMetGluAlaGly 520

Qy 2020 AGTACATGGGATTTTGTAACTCCACACCAACCAACGCTCTCTGCGCTAATCAGTTG 2079
Db 521 SerThrLeuAspPheValThrProGlnProGlnProAlaAlaAsnGlnLeu 540
Qy 2080 ATCAGCTTTTCCAATCTGCATTTGCTCTTCTTCTTTTGTAGCAACAATGCAGTTAGG 2139
Db 541 IleThrLeuSerAsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAlaValThr 560
Qy 2140 AATCTCTCTACCAATCTCCAGCGCAAGATCTCATCTCTGCACTCATTTGGTAGCAACT 2199
Db 561 AsnProProThrAsnProAlaGlnAspSerHisProAlaIleIleGlySerThrThr 580
Qy 2200 GCTGTTCTTGTTCACAAATAGTGGGCTATCTTTTGTAGGATTTGGATGATACAGCTTAT 2259
Db 581 AlaGlySerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyr 600
Qy 2260 GATAGTATGATGGCTAGTTCTTAATCAAAAATCAATGCTCTCAAAATACAGTTAGGG 2319
Db 601 AspArgTyrAspTyrLeuGlySerAsnGlnLysIleAspValLeuLysLeuGlnLeuGly 620
Qy 2320 ACTAAGCCCCCAGCTAATGCCCATCAGATTTGACTCTAGGGAATGAGATGCTAAGTAT 2379
Db 621 ThrGlnProSerAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyr 640
Qy 2380 GGCTATCAAGGAAGCTGGAAGCTTGTGGGATCCTTAATACAGCAAAATATGTCCTTAT 2439
Db 641 GlyTyrGlnGlySerTyrLysLeuAlaTyrPaspProAsnThrAlaAsnAsnGlyProTyr 660
Qy 2440 ACTCTGAAGCTACATGAGCTAAACTGGG 2469
Db 661 ThrLeuLysAlaThrTyrThrLysThrGly 670

RESULT 15
US-09-612-402B-17
; Sequence 17, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-17

Alignment Scores:
Pred. No.: 8,08e-228 Length: 505
Score: 2552.00 Matches: 505
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 32.4% Indels: 0
Gaps: 0
DB:

US-10-701-844-1 (1-4435) x US-09-612-402B-17 (1-505)
Qy 466 GAAATCATGTTCTCTCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTAT 525
Db 1 GluIleMetValProGlnGlyIleTyrAspGlyGluThrLeuThrValSerPheProTyr 20
Qy 526 ACTGTTATAGGAATCCGAGTGGGACTACTGTTTTTCTGCAGGAGTTAACTAAAA 585
Db 21 ThrValIleGlyAspProSerGlyThrValPheSerAlaGlyGluLeuThrLeuLys 40

141 SerPheaspGlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGly 160
1510 GCTATTATGCGAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGAAAT 1569
161 AlailetyrAlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuAgsen 180
1570 ATCGTAATCATGTGAGCGATTATTATTAGGAGATCTCGAGAGCTCAGTTTATCTGCT 1629
181 IleAlaAenaspGlyAlaAlailetyrLeuGlyGlySerGlyGlyLeuSerLeuSerAla 200
1630 GATTATGAGATATTATTTTCGATGGGAATCTTAAAGAAACAGCAAGAAAGATGCTGCC 1689
201 AspTyrglyAspIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAla 220
1690 GATGTTAATGCGTAATCTGCTCTCAAGCCATTTTCGATGGATCGGAGGGGAATA 1749
221 AspValAenGlyValThrValSerSerGlnAlaileSerMetGlySerGlyGlyLysile 240
1750 ACGACATTAAGCTAAAGCAGGCGCATCAGATTCTCTTAAATGATCCCATCGAGATGCCA 1809
241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAenaspProIleGluMetAla 260
1810 AACGGAATAACACGACGAGCGAGCTCTTCCAAACTTCTAAATAATTAACGATGGTGAAGGA 1869
261 AenGlyAenAenGlnProAlaGlnSerSerLysLeuLeuLysIleAenAspGlyGluGly 280
1870 TACACAGGGGATATGTTTGTCTAATGAAGAGAGTACTTGTACCAAAATGTTACGATA 1929
281 TyrThrGlyAspIleValPheAlaAenGlySerSerThrLeuTyrglnAenValThrIle 300
1930 GAGCAAGGAAGATGTTCTTCGTAAGCAAGCAAAATTTATCAGTGAATTTCTTAAGTCAG 1989
301 GluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSerValAenSerLeuSerGln 320
1990 ACAGTGGAGTCTGTATATGAAGCTGGGAGTACATGGGATTTTGTAACTCCCAACCA 2049
321 ThrGlySerLeuTyrglnMetGluAlaGlySerThrTrpAspPheValThrProGlnPro 340
2050 CCACAACAGCTCTCGCGCTAATCAGTTGATCAGCTTTCATCTGCATTTGCTCTT 2109
341 ProGlnGlnProProAlaAlaAenGlnLeuIleThrLeuSerAenLeuHisLeuSerLeu 360
2110 TCTCTTTGTAGCAACAATGAGTACAGTACAGTACCTCTTACCAATCTCCAGCGCAAGAT 2169
361 SerSerLeuLeuAlaAenAlaValThrAenProProThrAenProProAlaGlnasp 380
2170 TCTCATCTGCGATCATTTGGTAGCAACTGCTGCTTCTGTACAAATTTAGTGGGCTATC 2229
381 SerHisProAlaValIleGlySerThrThrAlaGlySerValThrIleSerGlyProIle 400
2230 TTTTGTAGGATTTGGATGATACAGCTTATGATAGTATGATGCTAGTGGTCTTCAATCAA 2289
401 PhePheGluAenLeuAenAspThrAlaTyrglyAspArgTyrglyLeuGlySerAenGln 420
2290 AAAATCAATGCTCGTAATACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 2349
421 LysIleAenValLeuLysLeuGlnLeuGlyThrLysProProAlaAenAlaProSerAsp 440
2350 TTGACTCTAGGAAATGAGATGCTAAGTATGCTATCAAGGAAGCTGGAAGCTT 2403
441 LeuThrLeuGlyAenGluMetProLysTyrglyTyrglnGlySerTrpLysLeu 458

RESULT 18
US-09-542-520-36
Sequence 36, Application US/09542520
Patent No. 6887843
GENERAL INFORMATION:
APPLICANT: Jackson, W. James
APPLICANT: Pace, John L.
TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
FILE REFERENCE: 7969-076-999
CURRENT APPLICATION NUMBER: US/09/542,520

CURRENT FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: PCT/US98/20737
PRIOR FILING DATE: 1998-10-01
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 36
LENGTH: 458
TYPE: PRT
ORGANISM: Chlamydia
US-09-542-520-36
Alignment Scores:
Pred. No.: 4,32e-209 Length: 458
Score: 2350.00 Matches: 458
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 29.8% Indels: 0
DB: Gaps: 0

US-10-701-844-1 (1-4435) x US-09-542-520-36 (1-458)
QY 1030 GGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGCTAACGAGGCTCTTATTGCC 1089
Db 1 GlycylAlaCysGlnValValThrSerPheSerAlaMetAlaAenGluAlaProIleAla 20
QY 1090 TTTGTAGCGAATGTTGAGGAGTAAGAGGGGGAGGATTCCTGCTGTTCAGGATGGGAG 1149
Db 21 PheValAlaAenValAlaGlyValArgGlyGlyIleAlaAlaValGlnAenGlyGln 40
QY 1150 CAGGAGTGTCATCATCTTCTTCAACAGAGATCCAGTAGTAACTTTTCCAGAAATCT 1209
Db 41 GlnGlyValSerSerThrSerThrGluAenProValValSerPheSerArgAenThr 60
QY 1210 GCGGTAGAGTTGATGGAGACGTAGCCGAGTAGGAGGAGGATTTTACTCTACGGGAAC 1269
Db 61 AlaValGluPheAspGlyAenValAlaArgValGlyGlyIleTyrglySerTyrglyAen 80
QY 1270 GTTCTCTTCTCGAATAATGGAACCACTTGTCTCAACAATGTTGCTTCTCTGTTTAC 1329
Db 81 ValAlaPheLeuAenAenGlyLysThrLeuPheLeuAenValAlaSerProValTyrg 100
QY 1330 ATTCTGCTTAAGCAACCAACAGTGGGAGGAGCTTCTTAACGAGTAAATTAACGGAGAT 1389
Db 101 IleAlaAlaLysGlnProThrSerGlyGlnAlaSerAenThrSerAenAenTyrglyAen 120
QY 1390 GAGGAGCTATCTTCTGTAAGATGGTGCGCAAGCAGGATCCAATACTCTGGATCAGTT 1449
Db 121 GlyGlyAlaIlePheCysLysAenGlyAlaGlnAlaGlySerAenAenSerGlySerVal 140
QY 1450 TCCTTTTATGAGAGGAGGAGTATTTTCTTTTAGTAGCAATGATGCTGCGGAAAGGGGA 1509
Db 141 SerPheAspGlyGluGlyValValPhePheSerSerAenValAlaAlaGlyLysGlyGly 160
QY 1510 GCTATTATGCGAAAGCTCTCGTTGCTTAACGTGGCCCTGTACAAATTTTAAAGAAAT 1569
Db 161 AlailetyrAlaLysLysLeuSerValAlaAenCysGlyProValGlnPheLeuAgsen 180
QY 1570 ATCGTAATCATGTGAGCGATTATTATTAGGAGATCTCGAGAGCTCAGTTTATCTGCT 1629
Db 181 IleAlaAenaspGlyAlaAlailetyrLeuGlyGlySerGlyGlyLeuSerLeuSerAla 200
QY 1630 GATTATGAGATATTATTTTCGATGGGAATCTTAAAGAAACAGCAAGAAAGATGCTGCC 1689
Db 201 AspTyrglyAspIlePheAspGlyAenLeuLysArgThrAlaLysGluAenAlaAla 220
QY 1690 GATGTTAATGCGTAATCTGCTCTCAAGCCATTTTCGATGGATCGGAGGGGAATA 1749
Db 221 AspValAenGlyValThrValSerSerGlnAlaileSerMetGlySerGlyGlyLysile 240
QY 1750 ACGACATTAAGCTAAAGCAGGCGCATCAGATTCTCTTAAATGATCCCATCGAGATGCCA 1809
Db 241 ThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAenaspProIleGluMetAla 260


```
/ APPLICANT: Pace, John L.
/ TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
/ FILE REFERENCE: 7969-076-999
/ CURRENT APPLICATION NUMBER: US/09/542,520
/ CURRENT FILING DATE: 2000-04-03
/ PRIOR APPLICATION NUMBER: PCT/US98/20737
/ PRIOR FILING DATE: 1998-10-01
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 325
/ TYPE: PRT
/ ORGANISM: Chlamydia
US-09-542-520-37

Alignment Scores:
Pred. No.: 4, 4e-152 Length: 325
Score: 1735.00 Matches: 325
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.0% Indels: 0
DB: 2 Gaps: 0

US-10-701-844-1 (1-4435) x US-09-542-520-37 (1-325)

QY 2443 CTGAAGCTACATGACATAAAGCTGGGTATATCTCGGCTGAGGAGTAGCTCTTTG 2502
Db 1 LeuysAlaThrTrpThrlysthrGlytyrAanProGlyProGluArgValAlaSerLeu 20
QY 2503 GTTCCAAATAGTTATGGGATCCATTATAGATATACGATCGCCATTCCAGCAATCAA 2562
Db 21 ValProAanSerLeuTrpGlySerIleuAaspIleAgsrAlaHisSerAlaIleGln 40
QY 2563 GCAAGTGTGGATGGCGCTCTTATGTCGAGGATATGAGTGTTCGAGTTTCGAATTC 2622
Db 41 AlaSerValAaspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAanPhe 60
QY 2623 TTCTATCATGACCGGATGCTTTAGGTACGGATATCGGTATATAGTGGGGGTATTCC 2682
Db 61 PheTyrHisAaspArgAaspAlaLeuGlyGlnGlytyrArgTyrIleSerGlytyrSer 80
QY 2683 TTAGGAGCAATCTCTACTTTGGATCATCGATGTTGGTCTAGCATTTACCAAGTATT 2742
Db 81 LeuGlyAlaAanSerTyrPheGlySerSerMetPheGlyLeuAlaPheThrGluValPhe 100
QY 2743 GGTAGATCTAAAGATTATGATGTGCTTCCATCATCATGCTTCATGATGATCCGTT 2802
Db 101 GlyArgSerLyaspTyrValValCysArgSerAanHisAlaCysIleGlySerVal 120
QY 2803 TATCTATCTACCAACAAGCTTTATGTGATCTTATTTGTCGGAGATCGGTTTATCCGT 2862
Db 121 TyrLeuSerThrGlnGlnAlaLeuCysGlySerTyrLeuPheGlyAaspAlaPheIleArg 140
QY 2863 GCTAGCTAGCGGTTGGGATCAGCATATGAACCTCATATACATATTCGACGAGGAGC 2922
Db 141 AlaSerTyrGlyPheGlyAanGlnHisMetLystrSerTyrThrPheAlaGluGluSer 160
QY 2923 GATGTTGCTGGGATTAATCACTGCTGCTGGCTGAGAGATTGGAGCGGATTACCGATTGTG 2982
Db 161 AspValArgTrpAaspAanCysLeuAlaGlyGluIleGlyValAaglyLeuProIleVal 180
QY 2983 ATTACTCCATCTAAGCTCTATTGAATGATGTCGTCCTTTGTCGAAGCTAGTTTCT 3042
Db 181 IleThrProSerLyLeuSerLeuAanGluLeuArgProPheValGlnAlaGluPheSer 200
QY 3043 TATGCCGATCATGAATCTTTTACAGAGGAGCGATCAAGCTCGGCATTTCAGAGCGGA 3102
Db 201 TyrAlaAaspHieGluSerPheThrGluGluGlyAaspGlnAlaArgAlaPheLysserGly 220
QY 3103 CATCTCTAAATCTATCATGTTCCGTTGGAGTGAAGTTTGATGCGATGTTCTAGTACACAT 3162
Db 221 HisLeuLeuAanLeuSerValProValGlyValLysePheAaspArgCysSerSerThrHis 240

/ APPLICANT: Steven P. Fling
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
/ TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
/ FILE REFERENCE: 210121.469C7
/ CURRENT APPLICATION NUMBER: US/09/620,412C
/ CURRENT FILING DATE: 2000-07-20
/ NUMBER OF SEQ ID NOS: 363
/ SOFTWARE: FastSeq for Windows Version 3.0/4.0
/ SEQ ID NO 325
/ LENGTH: 631
/ TYPE: PRT
/ ORGANISM: Chlamydia trachomatis
US-09-620-412C-325

Alignment Scores:
Pred. No.: 1, 68e-138 Length: 631
Score: 1590.50 Matches: 316
Percent Similarity: 88.7% Conservative: 7
Best Local Similarity: 86.8% Mismatches: 19
Query Match: 20.2% Indels: 22
DB: 2 Gaps: 4

US-10-701-844-1 (1-4435) x US-09-620-412C-325 (1-631)

QY 3391 TTGAGTCGAGGAGTAGAGTCGGTCTCTAAATAATAT-----TGGTAGATAGTTAA 3441
Db 129 LeuAlaGluGlyProProAlaGluPheCysArgTyrProSerHisTrpArgPro----- 146
QY 3442 GTGTTAGCGATGCCCTTTTCTTTGAGATCTACATCATTTGTTTGTAGCTTTGTGCT 3501
Db 147 -----LeuMetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCys 164
QY 3502 TCCTATTGCTATGGATTCCGAGCTCTCTCAAGTGTTAACGCCCTAATGTAAACCACTCT 3561
Db 165 SerTyrSerTyrGlyPheAlaSerSerProGlnValLeuThrProAanValThrThrPro 184
QY 3562 TTTAAGGAGGAGCATGTTTACTTGAATGGAGATCGCTTTGTCAATGTCTATGCAGGA 3621
Db 185 PheLyGlyAaspAaspValTyrLeuAanGlyAaspCysAlaPheValAanValTyrAlaGly 204
QY 3622 GCTGAAGAAGGTTTCCATTATCTCAGCTAATGGCGCAATTTAAACGATTACCGGACAAAAC 3681
Db 205 AlaGluAanGlySerIleIleSerAlaAanGlyAaspAanLeuThrIleThrGlyGlnAan 224
QY 3682 CATACATTATCATTTACAGATTCTCAAGGCGCATGTTCTTCAAAATATGCTTCATTCTCA 3741
Db 225 HisThrLeuSerPheThrAaspSerGlnGlyProValLeuGlnAanTyrAlaPheIleSer 244
```



```
QY 4306 GACAGGCGGTGGAGGAGTTATTTGCTCTCCAGATGATCTCTAAAGTTTGAAGGCAA 4365
D 442 UthrGlyGlyGlyGlyGlyValIleCysSerProAspAspSerValIlePheGluGlyAs 462
QY 4366 TAAAGGTTCTATTTGTTTGTATTAACAATTTTCAAAAGGCGAGGCGGAAGCATCTAAC 4425
D 462 nlysglySerIleValPheAspTyrAsnPheAlaIlysglyArgGlyGlySerIleLeuTh 482
QY 4426 GAAAGAAATTC 4435
D 482 rlysglyubhe 485

RESULT 23
US-10-197-220-95
; Sequence 95, Application US/10197220
; Patent No. 6919187
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 95
; LENGTH: 1016
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-10-197-220-95

Alignment Scores:
Pred. No.: 3,046-138 Length: 1016
Score: 1589.00 Matches: 311
Percent Similarity: 92.7% Conservative: 5
Best Local Similarity: 91.2% Mismatches: 10
Query Match: 20.2% Indels: 15
DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x US-10-197-220-95 (1-1016)

QY 3451 ARGCTTTTCTTGAGATCTACATCATTTGTTTGTAGTGTGTTGTTCCATTGCG 3510
D 1 MetProPheSerLeuArgSerThrSerPheCysPheLeuAlaCysLeuCysSerTyrSer 20
QY 3511 TATGATTTCGCGAGCTCTCTCAAGTGTAAAGCCCTTAATGTAACCACTCTTTTAAGGGA 3570
D 21 TyrGlyPheAlaSerSerProGlnValLeuThrProAsnValThrThrProPheLysGly 40
QY 3571 GACGATGTTTACTTGAATGGAGACTGCGCTTTTGTCAATGTTCTATGCGAGAGCTGAAGAA 3630
D 41 AspAspValTyrLeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsn 60
QY 3631 GGTTCGATTATCTAGCTAATCGGCAATTTAAAGATTACCGGACAAACCATCATTA 3690
D 61 GlySerIleIleSerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeu 80
QY 3691 TCATTATACGATCTCAAGGCGAGTCTTCAAAATTAATGCTTCAATTCACGAGAGAG 3750
D 81 SerPheThrAspSerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGlu 100
QY 3751 ACACCTACTCTCAGAGATTTTTCGAGTCTGATGTTCTCAAAATGTTTCTTCGCGAGAA 3810
D 101 ThrLeuThrLeuLysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlu 120
QY 3811 AAGGGAATGATCTCCGGGAAAACCGTAGTATTTCCGGAGCGAGGCGAAGTGAATTTCTGG 3870
D 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrp 140
QY 3871 GATACTCCGCGGGTATTTCTCTTTATCTATCTGTGCGCAACTCATCATCACTCGCCT 3930
D 121 LysGlyMetIleSerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheTrp 140
```

```
Db 141 AspAsnSerValGlyTyrSerProLeuSerIleValProAlaSerThrProThrProPro 160
QY 3931 GCT-----CCACAGTTAGTATGCTCGG 3954
D 161 AlaProAlaProAlaProAlaSerSerLeuSerProThrValSerAspAlaArg 180
QY 3955 AAGCGTCTATTTTCTGTAGAGACTAGTTTTCAGATCTCAGCGCTCAAAAAGGGGCTC 4014
D 181 LysGlySerIlePheSerValGluThrSerLeuGluIleSerGlyValLysGlyVal 200
QY 4015 ATGTTTCGATAATAATGCGGGAATTTTCGAAACAGTTTTTCGAGGTAAAGATAATAAT 4074
D 201 MetPheAspAsnAsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsn 220
QY 4075 GCTGGTGGTGGAGGAGTGGTTCGCTACACCATCAAGTACGACTTTTACAGTTAAAAA 4134
D 221 AlaGlySerGlyGlySer-GlySerAlaThrThrProSer-----PheThrValLysAs 238
QY 4135 CTGTAAAGGGAAGTTTCTTTTCACAGATAAGTAGCTCTTCGAGGCGGAGTGGTTTA 4194
D 238 nCysLysGlyLysValSerPheThrAspAsnValAlaSerCysGlyGlyGlyValVal 258
QY 4195 TAAAGGCATTGCTGCTTTTCAAGACAATGAAGGAGGCATATTTCTCCAGGGAACACAGC 4254
D 258 rlysglyThrValLeuPheLysAspAsnGluGlyLysIlePhePheArgGlyAsnThrAl 278
QY 4255 ATACGATGATTAAAGGATTTCTGCTGCTACTAATCAGCATCAGATCAGATCAGAGAGG 4314
D 278 aTyrAspAspLeuGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGl 298
QY 4315 CGGTGGAGGAGTATTTGCTCTCCAGATGATCTCTGTAAGTTTGAAGGCAATAAAGTTTC 4374
D 298 yGlyGlyGlyValIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySe 318
QY 4375 TATTGTTTTCATTACAACTTTTCAAAAGGCGAGGCGGAGCATCTTAACGAAAGAATT 4434
D 318 rIleValPheAspTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPh 338
QY 4435 C 4435
D 338 e 338

RESULT 24
US-10-197-220-168
; Sequence 168, Application US/10197220
; Patent No. 6919187
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Guderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 168
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Chlamydia
US-10-197-220-168

Alignment Scores:
Pred. No.: 5,3e-126 Length: 664
Score: 1456.00 Matches: 287
Percent Similarity: 92.1% Conservative: 5
Best Local Similarity: 90.5% Mismatches: 10
Query Match: 18.5% Indels: 15
DB: 2 Gaps: 2

US-10-701-844-1 (1-4435) x US-10-197-220-168 (1-664)
```

```

3523 AGCTCTCTCAAGTGTAGCGCTATATGTAACCACTCTTTTAAAGGAGACGATGTTTAC 3582
Db 1 SerSerProGlnValLeuThrProAsnValThrProPheLysGlyAspValTyr 20

3583 TTGAATGGAGACTGCCCTTTTGTCAATGCTATGACGAGAGCTGCAAGAGTTCGATATTC 3642
Db 21 LeuAsnGlyAspCysAlaPheValAsnValTyrAlaGlyAlaGluAsnGlySerIle 40

3643 TGAGCTAATGGGACAACTTACGATTACCGGACAAACCATACATATCATTTACAGAT 3702
Db 41 SerAlaAsnGlyAspAsnLeuThrIleThrGlyGlnAsnHisThrLeuSerPheThrAsp 60

3703 TCTCAAGGCGCAGTCTCTCAAAATATGCTTCATTTACAGCAGAGACACTTACTCTG 3762
Db 61 SerGlnGlyProValLeuGlnAsnTyrAlaPheIleSerAlaGlyGluThrLeuThrLeu 80

3763 AGAGATTTTTCGAGTCTGATGTTCTCGAAATATGTTCTTGGCGAGAAAGGAATGATC 3822
Db 81 LysAspPheSerSerLeuMetPheSerLysAsnValSerCysGlyGlyMetIle 100

3823 TCCGGGAAACCGTGAATTTCCGGAGACAGCGGAAGTATTTCTGGGATTAACCTCGTG 3882
Db 101 SerGlyLysThrValSerIleSerGlyAlaGlyGluValIlePheThrAspAsnSerVal 120

3883 GGGTATTCCTCTTATCTACTGTGCGCACTCATCATCACTCCGCTGCT----- 3933
Db 121 GlyTyrSerProLeuSerIleValProAlaSerThrProThrProProAlaProAlaPro 140

3934 -----CCAAACAGTTAGTATGCTCGGAAGGCTGATTT 3966
Db 141 AlaProAlaAlaSerSerSerLeuSerProThrValSerAspAlaArgLysGlySerIle 160

3967 TTTTCTGTAGACTAGTTGGAGATCTCAGCGCTCAAAAGGGGTGATGTTTCGATAT 4026
Db 161 PheSerValGluThrSerLeuGluIleSerGlyValLysGlyValMetPheAspAsn 180

4027 AATCGCGGNAATTCGGAACAGTTTTCGAGGTAAAGTAATTAATGCTGCTGGTGA 4086
Db 181 AsnAlaGlyAsnPheGlyThrValPheArgGlyAsnSerAsnAsnAlaGlySerGly 200

4087 GGCAGTGGGTTCGCTACACCATCAAGTACGACTTTTACAGTTTAAACCTGTAAAGGAA 4146
Db 201 GlySer-GlySerAlaThrThrProSer-----PheThrValLysAsnCysLysGly 218

4147 AGTTTCTTTCACAGATAACCTAGCTCTTCGCGAGCGGAGTGTGTTTATAAGGCATGT 4206
Db 218 sValSerPheThrAspAsnValAlaSerCysGlyGlyValValTyrLysGlyThrVa 238

4207 GCTTTTCAAGACAAATGAAGGCGCATATTTCTTCGAGGGAACACAGCATACGATGTTT 4266
Db 238 lLeuPheLysAspAsnGluGlyGlyIlePhePheArgGlyAsnThrAlaTyrAspAspLe 258

4267 AAGGATTTCTGCTACTAATCAGGATCAGATACGAGACAGGCGGCTGGAGAGT 4326
Db 258 uGlyIleLeuAlaAlaThrSerArgAspGlnAsnThrGluThrGlyGlyGlyGlyVa 278

4327 TATTTGCTCTCCAGATGATTTCTGTAAGTTTGAAGCAATAAAGTTTCTATTGTTTGA 4386
Db 278 lIleCysSerProAspAspSerValLysPheGluGlyAsnLysGlySerIleValPheAs 298

4387 TTACAACTTTCGAAAGCGAGCGGAGGAGCAAGCATCTAACGAAAGATTC 4435
Db 298 pTyrAsnPheAlaLysGlyArgGlyGlySerIleLeuThrLysGluPhe 314

```

RESULT 25

```

US-09-438-185A-455
; Sequence 455, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald

```

```

; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PR
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPN0453
US-09-438-185A-455

```

```

Alignment Scores:
Pred. No.: 1,37e-117 Length: 999
Score: 1366.50 Matches: 354
Percent Similarity: 49.7% Conservative: 177
Best Local Similarity: 33.1% Mismatches: 430
Query Match: 17.3% Indels: 107
DB: 2 Gaps: 28
US-10-701-844-1 (1-4435) x US-09-438-185A-455 (1-999)

```

```

QY 322 ATTAATAATTAATTTTATGAAGCGGAGTAATTAATTTTATCTCTC----- 369
Db 3 ValLeuLeuTyrLeuPhePheTyrSerLeuSerLeuIleCysArgIleIleTyrPheHis 22

QY 370 AGCTTTTGTGTGATGCAACGCTTTCCATAGAGTCTTCTTCTTCAATGATCTTAGCTTAT 429
Db 23 LeuTyrValGlnMetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaPro 42

QY 430 TCTTGCTGCTCTTTAAATGCGGGGGATATGCAGCAGAAATCATGTTCTCTCAAGGAATT 489
Db 43 CysPheAlaSerThr-----AlaPheThrValGluValIleMetProSerGluAsn 59

QY 490 TAGCATGGGAGGAGCTTAATCTGATCATTTCCCTACTGTTATATAGAGATCCGAGTGGG 549
Db 60 PheAspGlySerSerGlyLysIle---PheProTyrThrThrLeuSerAspProArgGly 78

QY 550 ACTACTGTTTCTGTCAGGAGAGTTAACAATTAATAAATCTTGACAATCTCTATTCGAGCT 609
Db 79 ThrLeuCysIlePheSerGlyAspLeuTyrIleAlaAsnLeuAspAsnAlaIleSerArg 98

QY 610 TTGCTTTAAGTTGTTTGGGAACTTATTAGGAGTTTACTGTTTTCGAGGAGACAC 669
Db 99 ThrSerSerSerCysPheSerAsnArgAlaGlyAlaLeuGlnIleLeuGlyLysGlyGly 118

QY 670 TCGTTCACCTTTCGAGAACATACGAGCTTCTACAAATGGGGGAGCTCTAAGTAATAGCGCT 729
Db 119 valPheSerPheLeuAsnIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIle 138

QY 730 GCTGATGGA-----CTGTTTACTATTGAGGTTTAAAGAAATATCTCTTTCC 777
Db 139 ThrGlnAsnProGluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAsp 158

QY 778 AATTGCAATTCATTACTTGGCGTACTGCTGCTGCAACGACTAATAAGGTCAGCAGACT 837
Db 159 AsnCysGluSerLeuThr-----SerAspThr 167

QY 838 CCGACGACACACTCTACACCGTCTAATGCTACTATTATTCTAAACACAGATCTTTTGTTA 897
Db 168 SerAlaSerAsnValIleProHisAlaSerAlaIleTyrAlaThrThrProMetLeuPhe 187

QY 898 CTCATAATGAAGTTCATTCTATAGTAATTTAGTCTCTCGAGATGGGAGCTATA 957
Db 188 ThrAsnAsnAspSerIleLeuPheGlnTyrAsnArgSerAlaGlyPheGlyAlaAlaIle 207

```



```

Db      872 GluValAlaTyAlaThrHisGlyGlyIleGlnGluAenThrProGluGlyArgIlePhe 891
QY      3094 AAGAGCGGAGCATCTCTTAATCTATCAGTTCTCTGTTGGAGTGAAGTTGATCGATGTTCT 3153
Db      892 GlyHisGlyHisLeuLeuAenValAlaValProValGlyValArgPheGlyLysAenSer 911
QY      3154 AGTACACATCTTAATAAATAGCTTTATGGCGGCTTATCTGTGATCTTATCGCACCC 3213
Db      912 HisAenArgProAspPheTyThrIleIleValAlaTyAlaProAspValTyArgHis 931
QY      3214 ATCTCTGGTACTGAGACAGCTCTATCCCATCAGACATGAGACATGAGACATGCGCTTT 3273
Db      932 AsnProAspCysAspThrThrLeuProIleAenGlyAlaThrTriphrSerIleGlyAen 951
QY      3274 CATTTAGCAAGATCGAGTTGTGGTTAGAGGATCTATGCTTCTCTCAACAGTAAT 3333
Db      952 AsnLeuThrArgSerThrLeuLeuValGlnAlaSerSerHisThrSerValAenAspVal 971
QY      3334 ATAGAAGTATATGGCCATCGAAGATATGATATCGAGATGCTTCTCGAGGCTATGTTTG 3393
Db      972 LeuGluIlePheGlyHisCysGlyCysAspIleArgThrSerArgGlnTyThrLeu 991
QY      3394 AGTCGAGGAAGTAGTCCGGTTC 3417
Db      992 AspIleGlySerIlyLeuArgPhe 999

RESULT 26
US-09-430-723-2
; Sequence 2, Application US/09430723
; Patent No. 6607730
; GENERAL INFORMATION:
; APPLICANT: Mordin, Andrew
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-011
; CURRENT APPLICATION NUMBER: US/09/430,723
; CURRENT FILING DATE: 1999-10-29
; EARLIER APPLICATION NUMBER: 60/106,590
; EARLIER FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/133,071
; EARLIER FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-430-723-2

Alignment Scores:
Pred. No.: 6e-117 Length: 973
Score: 1359.50 Matches: 349
Percent Similarity: 49.8% Conservative: 421
Best Local Similarity: 33.4% Mismatches: 103
Query Match: 17.2% Indels: 2
DB: 2 Gaps: 27

US-10-701-844-1 (1-4435) x US-09-430-723-2 (1-973)
QY      382 ATGCACAGCTTCCATAGTTCTTTTCAATGATTCAGCTTATCTTGTGCTCT 441
Db      1 MetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaProCysPheAlaSer 20
QY      442 TTAATGGGGGGGATATGACGACGAAATCATGTTCTCCTCAAGGAATTTACGATGGGAG 501
Db      21 Thr-----AlaPheThrValGluValIleMetProSerGluAenPheAspGlySer 37
QY      502 ACGTTAACTGTATCAATTTCCCTATGTTATAGAGATCGGAGTGGGACTACTGTTTTT 561
Db      38 SerGlyLysIle---PheProTyThrThrLeuSerAspProArgGlyThrLeuCysIle 56
QY      562 TCTGCAGAGGTTAAATTAATAAATCTTGACAAATCTATGCGCTTTGCGCTTTAAGT 621

```

```

Db      57 PheSerGlyAspLeuTyIleAlaAenLeuAspAenAlaIleSerArgThrSerSerSer 76
QY      622 TGTTTTGGAACTTATTAGGAGGATTTTACTGTTTGGGAGAGGACACTCGTTGATCTTTC 681
Db      77 CysPheSerAenArgAlaGlyAlaLeuGlnIleLeuGlyLysGlyValPheSerPhe 96
QY      682 GAGAACATACGAGCTTCTACAAATGGGCGAGCTCTAAATATAGCGCTGCTGATGGA--- 738
Db      97 LeuAenIleArgSerSerAlaAspGlyAlaAlaIleSerSerValIleThrGlnAenPro 116
QY      739 -----CTGTTTACTATTGAGGGTTTAAAGAAATTATCTTTTCAATTCGAATTC 789
Db      117 GluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAspAenCysGluSer 136
QY      790 TTTACTTGGCTACTGCTGCTGCAACGACTAATAAGGGTAGCGACACTCCGACGACAACA 849
Db      137 LeuThr-----SerAspThrSerAlaSerAen 145
QY      850 TCTACACCGTCTPAATGGTACTATTATTCTTAAACAGATCTTTTGTGTACTCAATATGAG 909
Db      146 ValIleProHisAlaSerAlaIleTyAlaThrThrProMetLeuPheThrAenAenApp 165
QY      910 AAGTTCTCATCTTATAGTAATTTAGTCTCTCGAGATGGGGAGCTATAGATGCTAAGAC 969
Db      166 SerIleLeuPheGlnTyArgSerAlaGlyPheGlyAlaAlaIleArgGlyThrSer 185
QY      970 TTAACGGTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAGAAATACTCTCAAGCTGAT 1029
Db      186 IleThrIleGluAenThrLysLysSerLeuLeuPheAenGlyAenGlySerIleSerAen 205
QY      1030 GGGGAGCTTGTCAAGTAGTACACAGTTTCTCTGCTATGCGCTAAACAGGCTCTTATGCC 1089
Db      206 GlyGlyAlaLeuThrGlySerAlaAlaIleAenLeuIleAenAenSerAlaProValIle 225
QY      1090 TTTGTAGCGAATGTTGCGAGGATAGAGGGGAGGAGGATTCGCTGTTCCAGGATGGCAG 1149
Db      226 PheSerThrAenAlaThrGlyIleTyGlyAlaIleTyLeuThr----- 241
QY      1150 CAGGAGTGTCTATCTACTTCAACAGAAAGTCCAGTAGTAGTATTTTCCAGAAATACT 1209
Db      242 ---GlyGlySerMetLeuThrSer-----GlyAenLeuSer 252
QY      1210 GCGGTAGAGTTTGTATGGGAACTGATCCCGAGTAGGAGGAGGATTTTACTCTACGGGAAC 1269
Db      253 GlyValLeuPheValAenAenSerSerArgGlyAlaIleTyAlaAenGlyAen 272
QY      1270 GTTGCTTCTCTGAATATGGAACCTTGTGTTTCTCAACAATGCTGCTCTCTGTT--- 1326
Db      273 ValThrPheSerAenAenSerAspLeuThrPheGlnAenAenThrAlaSerProGlnAen 292
QY      1327 TACATTGCTGCTAAGCAACCAACAGTGGACAGGCTTCTAATACGAGTAAATATTCGGA 1386
Db      293 SerLeuProAlaProThrProProThrProProAlaValThrPro---LeuLeuGly 311
QY      1387 GATGGAGGAGCTATCTTCTGTAAAGATGGTGCACAGCAGGATCCAATAACTCTGGATCA 1446
Db      312 TyrGlyGlyAlaIlePheCysThrProProAlaThrProProThrGlyVal---Ser 330
QY      1447 GTTTCCTTTGATGAGGAGGAGTGTGTTTCTTTTAGTACATAGTCTGCTGGGAAGGG 1506
Db      331 LeuThrIleSerGlyGluAenSerValThrPheLeuGluAenIleAlaSerGluGly 350
QY      1507 GAGACTTTATCCCAAAAGCTCTCGTTGCTTAACTGTGCTGCTGCTGCTGCTGCTGCTGCT 1566
Db      351 GlyAlaLeuTyGlyLysIleSerIleAspSerAenLysSerThrIlePheLeuGly 370
QY      1567 AATATCGCTAATGATGCTGGAGCGATTATTATTAGGAGAAATCTCGAGAGCTCAGTTTATCT 1626
Db      371 AsnThrAlaGlyLysGlyAlaIleAlaIleProGluSerGlyGluLeuSerLeuSer 390
QY      1627 GCTGATTATGAGATATTTATTTTCGATGGGAATCTTAAAGAACACAGCAAGAGAAATGCT 1686
Db      391 AlaAenGlnGlyAspIleLeuPheAenLysAenLeuSerIleThrSerGly----- 407

```


ORGANISM: Chlamydia pneumoniae
US-09-428-122-2

Alignment Scores:

Pred. No.: 1,46e-95 Length: 928
Score: 1129.00 Matches: 312
Percent Similarity: 45.7% Conservative: 171
Best Local Similarity: 29.5% Mismatches: 401
Query Match: 14.3% Indels: 174
DB: 2 Gaps: 30

US-10-701-844-1 (1-4435) x US-09-428-122-2 (1-928)

QY 382 ATGCAACGCTCTTTCATAGATCTCTTCTTCAATGATCTAGCTATTCTTCTGCTCT 441
DB 1 MetLysSerSerPheProLysPheValPheSerThrPheAlaIlePheProLysSerMet 20
QY 442 TTAATGGGGGGGATATGCAGAGAA---ATCATGGTTCTCTCAAGGAATTTACGATGG 498
DB 21 Ile-----AlaThrGluThrValLeuAspSerSerAlaSerPheAspGly 35
QY 499 GAGACGTTAACTGATATCATTTCCCTATCTGTTATAGGAGATCCGAGTGGGACTACTGTT 558
DB 36 AsnLys---AsnGlyAsnPheSerValArgGluSerGlnGluAspAlaGlyThrThrTyr 54
QY 559 TTTCTGCGAGAGATTACATTAATAAATCTTGCAATCTTATTCAGCTTTGCTTTA 618
DB 55 LeuPheLysGlyAsnValThrLeuGluAsnIleProGlyThrGlyThrAlaIleThrLys 74
QY 619 AGTTGTTTGGGAACATTATTAGGAGTCTTACTGTTTATAGGAGAGACACTCGTTGACT 678
DB 75 SerCysPheAsnAsnThrLysGlyAspLeuThrPheThrGlyAsnGlyAsnSerLeuLeu 94
QY 679 TTCGAGAACATCGGACTTCTACA---AATGGGGCAGCTCTAAGTAATAGCGCTGCTGAT 735
DB 95 PheGlnThrValAspAlaGlyThrValAlaGlyAlaAlaValAsnSerSerValValAsp 114
QY 736 GGACTGTTTACTATTGAGGGTTTAAAGATTTATCTCTTCCATTTGCAATTCATTACTT 795
DB 115 LysSerThrThrPheIleGlyPheSerSerLeuSerPhe-----IleAla 129
QY 796 GCGTACTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACATCTACA 855
DB 130 SerProGlySerSerIleThrThrGlyLysGlyAlaValSerCysSerThrGlySer--- 148
QY 856 CGCTCTAATGGTACTATTATTCTAAACACAGACTTTTGTGTTACTCAATAATGAGAAGTTC 915
DB 149 -----LeuSerLeuThrLysAsnValSerLeu 157
QY 916 TCATTCATAGTAATTTAGTCTCTGGAGATGGGGAGCTATAGATGCTAAGACTTAACG 975
DB 158 LeuPheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrLeuSer 177
QY 976 GTTCAGAGGATTAGCAAGCTTTGTGTTCTCCAGAAAATACTCTCAAGCTGATGGGGA 1035
DB 178 LeuThrGlyThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGlyGly 197
QY 1036 GCTTGTCAAGTAGTACACAGTTTCTCTGCTATGGCTAAAGAGGCTCTATTTGCTTTGTA 1095
DB 198 AlaIleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPhe--- 216
QY 1096 GCGAATGTCAGAGATGAGAGGGGGAGGATGCTCTGTTCCAGATGGGACAGGGA 1155
DB 216 ----- 216
QY 1156 GTGTCATCATCTACTTCAACAGAGATCCAGTAGTAGTTTTCAGAAATAGTCTGGGTA 1215
DB 217 ---SerAspAsnThrSerSerAspSer----- 224
QY 1216 GAGTTTGTGGAACGATAGCCGAGTAGGAGGAGGATTTACTCTACGGGAACGTTGCT 1275
DB 225 -----GlyAlaAlaIlePheThrGluAlaSerValThr 235

QY 1276 TTCGTAATAATGCAAAACCTTGTCTCAACAATGTTGCTTCTCTCTGTTTACATTCCT 1335
DB 236 IleSerAsnAsnAlaLysValSerPheIleAspAsn----- 247
QY 1336 GCTAAGCAACCAACAGTGGAGCAGCTTCTAATAGCAGTAATAATTACGAGATCGAGA 1395
DB 248 -----LysValThrGlyAlaSerSerThrThrGlyAspMet---SerGlyGly 263
QY 1396 GCTATCTTCTGTAAGATGGTGGCGAAGCAGATCCAATAACTCTGGATCAGTTTCCCTTT 1455
DB 264 AlaIleCysAlaTyrLys-----ThrSerThrAspThrLysValThrLeu 278
QY 1456 GATGAGAGGAGTAGTATTTCTTTAGTAGCAATGATGCTGGGAAAGGGGAGGATATT 1515
DB 279 ThrGlyAsnGlnMetLeuPheSerAsnAsnThrSerThrAlaGlyAlaIle 298
QY 1516 TATGCCAAAAGCTCTCGTTCCTAATCTGGCCCTGTACAATTTTAAAGGAATATCGCT 1575
DB 299 TyrValLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerVal 318
QY 1576 AAT-----GATGGTGGAGCGATTATTATTAGGAGAAATCTGGAGAGCTCAGT 1620
DB 319 AsnGlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGluLeuSer 338
QY 1621 TTATCTGCTGATTTATGAGATATTATTTTCGATGGGAATCTTTAAAGAACACACAAAGAG 1680
DB 339 LeuSerAlaAspSerGlyAspIleValPheLeuGlyAsn-----ThrValThrSer 355
QY 1681 AATGCTGGGATGTTAATGGCGTAATCTGTCTCTCAAGCCATTTTCGATGGATCGGGA 1740
DB 356 ThrThrProGlyThrAsn-----ArgSerSerIleAspLeuGlyThrSer 370
QY 1741 GGGAAATAACACATTAAGACTAAAGCAGGCGATCATGATTTCTTTTAAATGATCCATC 1800
DB 371 AlaLysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleThrPheTyrAspProIle 390
QY 1801 GAGATGGCAACCGAAATAAACAGCCAGCGCAGTCTTCCAAACTTCTTAAATAATACGAT 1860
DB 391 ThrThrGlySerSerThrThr-----ValThrAspValLeuLysValAsnGlu 406
QY 1861 GTGGAAGGA-----TACACAGGGATATTGTTTTCCT----- 1893
DB 407 ThrProAlaAspSerAlaLeuGlnThrGlyAsnIlePheThrGlyGluLysLeu 426
QY 1894 -----AATGGAAGCAGTACTTTGTACCAAAATGTTTACG 1926
DB 427 SerGluThrGluAlaAlaAspSerLysAsnLeuThrSerLysLeuGlnProValThr 446
QY 1927 ATAGAGCAAGGAGGATTTCTTCTGTAAGGCAAAATTTATCAGTGAATTCCTTAAGT 1986
DB 447 LeuSerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThrGlnAlaPheThr 466
QY 1987 CAGACAGGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGGATTTGTAACCTCA 2043
DB 467 GlnGlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu----- 482
QY 2044 CAACACCAACACAGCTCTCTCGCTAATCAGTTTCATCAGCTTCCATCTGCAATTCG 2103
DB 483 -----ProAlaAspThrSer-----ThrIleAsnLeuValIle 494
QY 2104 TCTCTTTCTTCTTGTAGCAAAATGAGTTTACGAATCTCTCTACCAATCTCTCCAGCG 2163
DB 495 AsnIleSerSerIle----- 499
QY 2164 CRAGATTCTCATCTCCAGTCAATGTTAGC---ACAATGCTGCTGTTCTGTTACAATTAGT 2220
DB 500 AspGlyAlaLysAlaLysIleGluThrLysAlaThrSerLysAsnLeuThrLeuSer 519
QY 2221 GGCCTATCTTTTTCAGGATTTGGATGATACAGCTTATCATAGTAGTATGTTGCTAGGT 2280
DB 520 GlyThrIleThrLeuLeuAspProThrGlyThrPheTyrGluAsnHisSer---LeuArg 538
QY 2281 TCTAATCAAAAAATCAATGTCCTGMAATTACAGTTA---GGGACTAAGACCCCGAGCTAAT 2337

```
Db 539 AsnProGlnSerTyrAspIleLeuGluLeuIysAlaSerGlyThrValThrSerThrAla 558
QY 2338 GCGCCATCAGATTTCGACTCTAGGGAATCAGATGCCCTAAGTATGGCTATCAAGGAGCTGG 2397
Db 559 ValThrProAspProIleMetGlyGluIysPhe---HisTyrGlyTyrGlnGlyThrTrp 577
QY 2398 ---AAGCTTGCGTGGATCCTTAATACAGCAATAATAGTCTCTTATATCTCTGAAGCTACA 2454
Db 578 GlyProIleValTrpGlyThrGlyAlaSerThrThrAlaThrPhe-----Asn 593
QY 2455 TGGACTAAACCTGGGTATAAATCCTGGGCTGAGGAGTAGCTCTCTTGGTTCCTCAATAGT 2514
Db 594 TrpThrIysThrGlyTyrIleProAsnProGluArgIleGlySerLeuValProAsnSer 613
QY 2515 TTATGGGATCCATTTAGATATACGATCTGGCATTTCAGCAATTCAGCAAGTGTGTGAT 2574
Db 614 LeuTrpAsnAlaPheIleAspIleSerSerLeuHisTyrLeuMetGluThrAlaAsnGlu 633
QY 2575 GGGCGCTCTTATGTGCGAGGATTATGGGTTCTCGAGTTTCGAGTTTCGAAATTCCTTATCATGAC 2634
Db 634 GlyLeuGlnGlyAspArgAlaPheTrpCysAlaGlyLeuSerAsnPheHisIysAsp 653
QY 2635 CGCGATGCTTAGTTCAGGATATCGGTATATAGTGGGGTTATCTTCTAGGACCAAC 2694
Db 654 SerThrIysThrArgArgGlyPheArgHisLeuSerGlyIleTyrValIleGlyAsn 673
QY 2695 TCCTACTTTGGATCA---TCGATGTTGCTAGCAATTCACGAAATTTGGTAGACT 2751
Db 674 LeuHisThrCysSerAspIysIleLeuSerAlaAlaPheCysGlnLeuPheGlyArgAsp 693
QY 2752 AAAGATTATGATGTGTCGTTCCATCATCATGCTTCGATAGGATCCGTTTATCTATCT 2811
Db 694 ArgAspTyrPheValAlaIysAsnGlnGlyThrValTyrGlyIleThrLeuTyr-Tyr-Gl 713
QY 2812 ACCCAACAGCTTTATGTCGATCCCTATTTGTCGGAGATGCGTTTATCGTGCT----- 2865
Db 713 nHisAsnGluThrTyrIleSerLeuProCysIysLeuArgProCysSerLeuSerTyrVa 733
QY 2866 -----AGCTACGGGTTTGGAAATCAGCA 2888
Db 733 LProThrGluIleProValLeuPheSerGlyAsnLeuSerTyrThrHisThrAspAsnAs 753
QY 2889 TATGAAAACCTCATATACATTTGACAGAGCGATGTTCTGGGATAATAACTGTCT 2948
Db 753 pleuIysThrIysTyrThrTyrTrpProThrValIysGlySerTrpGlyAsnAspSerPh 773
QY 2949 GCGTCGAGATGTCAGCGGATACCGATTGTGATTACTCCTAAGCTCTATTGAA 3008
Db 773 eAlaLeuGluPheGlyIleArgAlaProIleCysLeuAspGluSerAlaLeuPhe---Gl 792
QY 3009 TGAGTTGCGTCTTTCGTCAGCTGAGTTTCTTATGCGGATCATGAATCTTTTACAGA 3068
Db 792 uGlnTyrMetProPheMetLeuGlnPheValTyrAlaHisGlnGluGlyPheIysGlu 812
QY 3069 GGAAGCGCATCAAGCTCGGGCATTCAGAGCGGACATCTCTAAATCTATCATCTCTGT 3128
Db 812 uGlnGlyThrGluAlaArgGluPheGlySerSerArgLeuValAsnLeuAlaLeuProIle 832
QY 3129 TCGAGTGAAGTTTGATCGATGTTCTAGTACATCATCTAATAATATAGCTTTATGCGGC 3188
Db 832 eGlyIleArgPheAspLysGluSerAspCysGlnAspAlaThrTyrAsnLeuThrLeuGl 852
QY 3189 TTATATCTGTGATGCTTATCGACCATCTCTGTACTGAGACAACGCTCTCTATCCATCA 3248
Db 852 yTyrThrValAspLeuValArgSerAsnProAspCysThrThrThrLeuArgIleSerGl 872
QY 3249 AGAGACATGACACAGATGTCCTTTTCATTAGCAAGACATGAGGTGTGCTTAGGATC 3308
Db 872 yAspSerTyrThrPheGlyThrAsnLeuAlaArgGlnAlaLeuValLeuArgAlaGl 892
QY 3309 TATGATGCTTCTCTAACAGTAATATAGAAATATAGCCATGCGGATATAGTATCG 3368
Db 3368
```

```
Db 892 yAsnHisPheCysPheAsnSerAsnPheGluAlaPheSerGlnPheSerPheGluLeuAr 912
QY 3369 AGATGCTTCTCGAGGCTATGTTTGTAGTGCAGGAGTAGAGTCCGGTTC 3417
Db 912 gGlySerSerArgAsnTyrAsnValAspLeuGlyAlaIysTyrGlnPhe 928
RESULT 28
US-09-438-185A-453
; Sequence 453, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 453
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURES:
; OTHER INFORMATION: Cpn0451
US-09-438-185A-453
Alignment Scores:
Pred. No.: 2,39e-92 Length: 967
Score: 1094.50 Matches: 320
Percent Similarity: 42.6% Conservative: 157
Best Local Similarity: 28.6% Mismatches: 414
Query Match: 13.9% Indels: 229
Gaps: 32
DB:
US-10-701-844-1 (1-4435) x US-09-438-185A-453 (1-967)
QY 286 AGGACGAATCGCTGCAAGATAAGA-----ACATTTATTCATATT 324
Db 1 LysThrGlnArgValIysIleLeuAspSerCysPheValIlePheAsnLeuIle 20
QY 325 AAATTTAATTTTATGAAGCGGAGTAATTAATTTATCTCTCAGCTTTGTGTGATG 384
Db 21 TyrLeuPheCysPheTyrIleAspAlaAsnSerSerLeuLysAsnLysSerIleThrMet 40
QY 385 CAAACGCTTCTCCATAAGTCTTCTTCAATGATTCCTAGCTTATCTTTCG-----TGC 438
Db 41 LysThrSerIleProTrpValLeuValSerSerValLeuAlaPheSerCysHisLeuGln 60
QY 439 TCTTTAAATGGGGGGGATATGCAGAGAAATCATGGTTCCTCAAGGAATTTACGATGG 498
Db 61 SerLeuAlaAsn-----GluGluLeuLeuSerProAspAspSerPheAsnGly 76
QY 499 GAGCGTTAATGATTCATTT-----CCCTATATCTGTTATAGA 537
Db 77 AsnIleAspSerGlyThrPheThrProLysThrSerAlaThrThrTyrSerLeuThrGly 96
QY 538 GATCCGAGTGGGACTACTGTTTCTCGCAGGAGAGTAAACATTAAATAAATCTTGACAA 597
Db 97 Asp-----ValPhePheTyrGluProGlyLysGlyThr----- 107
QY 598 TCTATTGCGACTTTCCTTTA-----AGTTGTTTGGGAACCTTATAGGAGTTTACT 651
Db 108 -----ProLeuSerAspSerCysPheLysGlnThrThrAspAsnLeuThr 122
QY 652 GTTTTAGGAGGAGGACACTCGTTGACTTCGAGACATACGAGCTTCTACAAATGGGCA 711
```

123	Db	PheLeuGlyAsnGlyHisSerLeuThrPheGlyPheIleAspAlaGlyThrHisAlaGly	142
712	Qy	GCTCTAAGTAATAGCGCTGCTGATGGACTGTTTACTATTGAGGGTTTTAAAGAAATATCC	771
143	Db	AlaAlaAlaSerThrThrAlaAsnIleAsnLeuThrPheSerGlyPheSerIleuLeuSer	162
772	Qy	TTTTCCAAATGCAATTCATTACTTGCCTACTGCCTGCTGCACGACTAATAAGGGTAGC	831
163	Db	PheAspSer-----	165
832	Qy	CAGACTCCGACGACAACATCTACACCGTCTAAATGGTACTATTTATCTTAAACAGATCTT	891
166	Db	---SerProSerThrThrValThrThrGlyGlnGlyThrLeuSerSerAlaGlyGlyVal	184
892	Qy	TTGTACTCAATAATGAGAGATTCTCATCTATAGTAATTTAGTCTCTCGAGATGGGGA	951
185	Db	AsnLeuGluAsnIleAerGlyLeuValAlaGlyAsnPheSerThrAlaAspGlyGly	204
952	Qy	GCTATAGATGCTAAGAGCTTAAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAA	1011
205	Db	AlaIleLeuGlyAlaSerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSerAsn	224
1012	Qy	AATACTGCTCAAGCTGATGGGGAGCTGTGCAAGTAGTACACGATTTCTCTGCTATGGCT	1071
225	Db	AsnSerSerSerThrLysGlyAla-----	233
1072	Qy	AACGAGGCTCTATTGCGCTTTGTAGCGCAATGTTGACGAGCTAAGAGGGGAGGATGCT	1131
234	Db	-----IleAlaThrThrAlaGlyAlaArg-----	241
1132	Qy	GCTGTTCCAGGATGGGACGACGGAGTGTCTATCATCTTCAACAGAGAATCCAGTAGTA	1191
241	Db	-----	241
1192	Qy	AGTTTTCAGAAATACTGCG---GTAGAGTTTGAATGGGAACGTAGCCGAGTAGGAGA	1248
242	Db	---IleAlaAsnAsnThrGlyTyValArgPheLeuSerAsnIleAlaSerThrSerGly	260
1249	Qy	GGGATTACTCTCAGCGGAAGTTCCTCTCGAATAATGCGAAAACCTTGTTCTCAAC	1308
261	Db	GlyAlaIleAspAspGluGlyThrSerIleuSerAsnAsnLysPheLeuTyPhe---	279
1309	Qy	AATGTTGCTTCTCCTGTTTACATTGCTGCTAAGCAACCAAGTGGACAGGCTTCTAAT	1368
280	Db	-----GluGlyAsnAlaAlaLys	285
1369	Qy	ACGAGTAATAATACGAGATGAGGAGCTATCTCTGTGAAGATGCTGCAAGCAGGA	1428
286	Db	ThrThr-----GlyGlyAlaIle---CysAsnThrLysAlaSerGlySer	299
1429	Qy	-----TCCAATAACTCGATCAGTTTCCTTTGATGAGAGGAGGATGTT	1473
300	Db	ProGluLeuIleLeuSerAsnAsnLys-----ThrLeu	310
1474	Qy	TTCCTTAGTAGCAATGATGCTGCTGGAAAAGGGGAGCTATTTATGCCAAAAAGCTCTCG	1533
311	Db	IlePheAlaSerAsnValAlaGluThrSerGlyGlyAlaIleHisAlaLysIleuAla	330
1534	Qy	GTGTCTAATCTGGCCCTGTACAAATTTTAAAGGAATATCGCTAATGAT---	1581
331	Db	LeuSerSerGlyGlyPheThrGluPheLeuArgAsnAsnValSerSerAlaThrProLys	350
1582	Qy	GCTGAGCGGATTTATTTAGGAGAAATCTCGAGAGCTCAGTTTATCTGCTGATTTAGGAGAT	1641
351	Db	GlyGlyAlaIleSerIleAspAlaSerGlyGluLeuSerLeuSerAlaGluThrGlyAsn	370
1642	Qy	ATTATTTTCGATGGGAATCTTTAAAGAACACGCCAAGAGAAATGCTGCCGATGTTAATGCG	1701
371	Db	IleThrPheValArgAsnThrLeuThrThr-----GlySerThrAsp-----	385
1702	Qy	GTAACGTGTCTCTCAAGCCATTTTCGATGGGATCGGAGGGGAAAAATAACGACATTAAGA	1761

386	Db	---ThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGluLeuArg	404
1762	QY	GCTAAAGCAGGGCATCAGATTCTCTTTAAATGATCCCATGAGATGCAAAACGGAATAAC	1821
405	Db	AlaAlaLysAsnHisThrIlePhePheTyrAspProIle	418
1822	QY	CAGCAGCGCAGTCTCCAAACTCTAAAATTAACAGATGGTGAAGGA	1869
419	Db	SerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeuAsn	438
1870	QY	---TACACAGGGATATTGTTTTTGGT	1893
439	Db	ProTyrGlnGlyThrIleLeuPheSerGlyLysThrLeuThrAlaAspGluLeuLysVal	458
1894	QY	---AATGGAAGCAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGAGATGTT	1947
459	Db	AlaAspAsnLeuLysSerPheThrGlnProValSerLeuSerGlyGlyLysLeu	478
1948	QY	CTTCGTCAAAAGCAAAATTCAGTGAATCTCTAAGTCAGACAGGTGGAGCTCTGAT	2007
479	Db	LeuGlnLysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeuLeu	498
2008	QY	---ATGGAAGCTGGGAGTACATGGGATTTGTAACTCCACCAACCAACAGCCCTCT	2064
499	Db	GlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySer	512
2065	QY	GCCGCTAATCAGTTTGATCAGCTTTCOAATCTGCATTTGTCTCTTCTTCTTTGTAGCA	2124
513	Db	---IleThrIleThrAsnLeuGlyIleAsnValAspSerLeuGlyLeu	527
2125	QY	AACAATGCAGTTAGCATCTCTCTACCAATCCTCCAGCGCAAGATTCTCATCTCGCAGTC	2184
528	Db	LysGlnProVal	533
2185	QY	ATTGCTAGCACAACTGCTGGTTCGTGTACAATTAGTGGGCTATCTTTTGGAGATTG	2244
534	Db	ThrAlaLysGlyAlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAspIle	553
2245	QY	GATGATACAGCTTATGATAGTATGATGTGGCTAGGTTCTAATCAAAAATCAATGTCCTG	2304
554	Db	GluGlyAsnIleTyrGluSerHis---MetPheSerHisAspGlnLeuPheSerLeuLeu	572
2305	QY	AAATTACAGTTAGGAGCTAAGCCCCAGCTAATGCCCCCATCAGATTGTGACTTAGGGAAT	2364
573	Db	LysIleThrVal	588
2365	QY	GAGATGCTCT	2400
589	Db	LeuIleProValProAlaGluAspProAsnSerGluTyrGlyPheGlnGlyGlnTrpAsn	608
2401	QY	CTTCGGTGGGATCTTAATACAGCAAAATAATGGTCTTATCTCTGAAAGCTACATGGACT	2460
609	Db	ValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrpThr	627
2461	QY	AAACTGGGTATATCTGGGCCCTGAGGAGTACTCTTTGGTTCCAAATAGTTTATGG	2520
628	Db	LysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeuTrp	647
2521	QY	GGATCCATTTTAGATATACGATCTCGCATTCAGCAATTCAAGCAAGTGTGGATGGCGC	2580
648	Db	GlyValPheThrAspIleArgSerLeuGlnGlnLeuValGluIleGlyAlaThrGlyMet	667
2581	QY	TCTTATTCGAGGATATAGGTTCTCGAGATTTTCGAATTTCTTCTATCATGACCGCAT	2640
668	Db	GluHisGlnGlnPheTrpValSerSerMetThrAsnPheLeuHisLysThrGlyAsp	687
2641	QY	GCTTTAGTCAGGATATCGGTATATATAGTCGGGGTTATTCTTTAGGAGCAATCTCTAC	2700
688	Db	GluAsnArgLysGlyPheArgHisThrSerGlyGlyTrpValIleGlyGlySerAlaHis	707
2701	QY	TTT---GGATCATCATGTTTGGTCTAGCATTTTACCGAAGTATTTTGGTAGATCTAAAGAT	2757
708	Db	ThrProLysAspAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLysAsp	727

1246 GGAGGGATTACTCTACGGGAACGTTGCTTTCTGAATAATGGAAGAAACCTGTTTCTC 1305
Db ||||| : : : : : ||||| : : : : :
242 GlyGlyAlaIleAspGluGlyThrSerIleLeuSerAsnAsnLysPheLeuYrPhe 261
Qy ||||| : : : : : ||||| : : : : :
1306 AACAAATGTTCTCTCTCTGTTTACATTGCTGCTAAAGCAACCAAGTGGAGAGCTTCT 1365
Qy ||||| : : : : : ||||| : : : : :
262 -----GluGlyAsnAlaIle 266
Db ||||| : : : : :
1366 AATACGAGTAATAATTACGGAGATGGAGGAGTATCTTCTGTAAGATGGTGGCAAGCA 1425
Qy ||||| : : : : : ||||| : : : : :
267 LysThrThr-----GlyGlyAlaIle---CysAsnThrLysAlaSerGly 280
Db ||||| : : : : : ||||| : : : : :
1426 GGA-----TCCAATAACTCTGGATCAGTTTCTCTTGATGGAGAGGAGTA 1470
Qy ||||| : : : : : ||||| : : : : :
281 SerProGluLeuIleSerAsnLys-----Thr 291
Db ||||| : : : : : ||||| : : : : :
1471 GTTTTCTTTAGPAGCAATGAGTCTGCTGGAAAGGGAGCTATTATTATGCCAAAAGCTC 1530
Qy ||||| : : : : : ||||| : : : : :
292 LeuIlePheAlaSerAsnValAlaGluThrSerGlyAlaIleHisAlaLysLysLeu 311
Db ||||| : : : : : ||||| : : : : :
1531 TCGGTTGCTAACTGTGGCCCTGTACATATTTTAAAGAAATATCGTAATGAT----- 1581
Qy ||||| : : : : : ||||| : : : : :
312 AlaLeuSerSerGlyGlyPheThrGluPheLeuArgAsnAsnValSerSerAlaThrPro 331
Db ||||| : : : : : ||||| : : : : :
1582 ---GGTGGAGCATTTATTAGGAGAACTCTGAGAGCTCAGTTTATCTGCTGATTATGA 1638
Qy ||||| : : : : : ||||| : : : : :
332 LysGlyAlaIleSerIleAspAlaSerGlyGlyLeuSerLeuSerAlaGluThrGly 351
Db ||||| : : : : : ||||| : : : : :
1639 GATATTATTTCGATGGGAATCTTAAAGAACAGCCAAAGAAATGCTCCGATGTTAAT 1698
Qy ||||| : : : : : ||||| : : : : :
352 AsnIleThrPheValArgAsnThrLeuThrThr-----GlySerThrAsp----- 367
Db ||||| : : : : : ||||| : : : : :
1699 GCGGTAACGTGCTCTCAACAGCAATTCATGGAGATCGGGAGGGAGAAATAAGCATTA 1758
Qy ||||| : : : : : ||||| : : : : :
368 -----ThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGluLeu 385
Db ||||| : : : : : ||||| : : : : :
1759 AGAGCTAAAGCAGGATCAGATTCTCTTAAATGATCCCATCGAGATGCCAAAGCAAT 1818
Qy ||||| : : : : : ||||| : : : : :
386 ArgAlaAlaLysAsnHisThrIlePhePheThrAspProIle----- 399
Db ||||| : : : : : ||||| : : : : :
1819 AACACGACCGCAGCTTCTCCAAACTTCTAAAAATTAAACGATGGTGAAGGA----- 1869
Qy ||||| : : : : : ||||| : : : : :
400 ThrSerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeu 419
Db ||||| : : : : : ||||| : : : : :
1870 -----TACACAGGGATATTGTTTGTCT----- 1893
Qy ||||| : : : : : ||||| : : : : :
420 AsnProTyrglnGlyThrIleLeuPheSerGlyGluThrLeuThrAlaAspGluLeuLys 439
Db ||||| : : : : : ||||| : : : : :
1894 -----AATGGAAGCAGTACTTGTACCAAAATGTTACGATAGCAGCAAGCAAGATT 1944
Qy ||||| : : : : : ||||| : : : : :
440 ValAlaAspAsnLeuLysSerPheThrGlnProValSerLeuSerGlyGlyLysLeu 459
Db ||||| : : : : : ||||| : : : : :
1945 GTTCTTCGTGMAAGCAAAATATCAGTGAATCTCTAAGTCAGACAGGTGGAGTCTG 2004
Qy ||||| : : : : : ||||| : : : : :
460 LeuLeuGlnLysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeu 479
Db ||||| : : : : : ||||| : : : : :
2005 TAT---ATGGAAGCTCGGAGTACATGGGATTTTGTAACTCCACCAACCAACAGCCT 2061
Qy ||||| : : : : : ||||| : : : : :
480 LeuGlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySer----- 494
Db ||||| : : : : : ||||| : : : : :
2062 CTGCGCGCTAATCAGTTGATCAGCTTTCCAATCTGCATTTGTCTCTTCTCTTTGTTA 2121
Qy ||||| : : : : : ||||| : : : : :
495 -----IleThrIleThrAsnLeuGlyIleAsnValAspSerLeuGly 508
Db ||||| : : : : : ||||| : : : : :
2122 GCAACAAATGCAGTTACGATCTCTACCAATCTCCAGCGCAAGATTCTCATCTGCA 2181
Qy ||||| : : : : : ||||| : : : : :
509 LeuLysGlnProVal-----Ser 514
Db ||||| : : : : : ||||| : : : : :
2182 GTCATTGGTAGCAACAACCTGCTGTTCTGTATCAAAATTAGTGGGCTATCTTTTGGAGAT 2241
Qy ||||| : : : : : ||||| : : : : :
515 LeuThrAlaLysGlyAlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAsp 534
Db ||||| : : : : : ||||| : : : : :
2242 TTGGATGATACAGCTTATGATAGTATGATGGCTAGGTCTTAATCAAAAAATCAATGTC 2301
Qy ||||| : : : : : ||||| : : : : :
535 IleGluGlyAsnIleTyrgluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553
Qy ||||| : : : : : ||||| : : : : :
2302 CTGAAATATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTGTACTTAGGG 2361
Db ||||| : : : : : ||||| : : : : :
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569
Qy ||||| : : : : : ||||| : : : : :
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397
Db ||||| : : : : : ||||| : : : : :
570 SerLeuIleProValProAlaAspProAsnSerGluTyrglyPheGlnGlnTrp 589
Qy ||||| : : : : : ||||| : : : : :
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457
Db ||||| : : : : : ||||| : : : : :
590 AsnValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrp 608
Qy ||||| : : : : : ||||| : : : : :
2458 ACTAAAACTGGTATAATCTCGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517
Db ||||| : : : : : ||||| : : : : :
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628
Qy ||||| : : : : : ||||| : : : : :
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577
Db ||||| : : : : : ||||| : : : : :
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648
Qy ||||| : : : : : ||||| : : : : :
2578 CGCTCTTATGTCGAGGATATGCGTTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637
Db ||||| : : : : : ||||| : : : : :
649 MetGluHisLysGlnGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 668
Qy ||||| : : : : : ||||| : : : : :
2638 GATGCTTATGTCGAGGATATCGGTATATAGTGGGGTTATCTCTAGGAGCAAACTCC 2697
Db ||||| : : : : : ||||| : : : : :
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyTyrglyValIleGlySerAla 688
Qy ||||| : : : : : ||||| : : : : :
2698 TACTTT---GGATCATCGATGTTTCTAGCATTTTACCGAAGTATTTGTTAGATCTAAA 2754
Db ||||| : : : : : ||||| : : : : :
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708
Qy ||||| : : : : : ||||| : : : : :
2755 GATTATGATGTCGTCTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814
Db ||||| : : : : : ||||| : : : : :
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrglyGlyThrLeuPhePheLysHis 728
Qy ||||| : : : : : ||||| : : : : :
2815 CAACAGCTTTA-----TGTGGATCTCTTTGTTGGAGATGG 2853
Db ||||| : : : : : ||||| : : : : :
729 SerHisThrLeuGlnProGlnAsnTyrglyLeuArgLeuGlyArgAlaLysPheSerGluSer 748
Qy ||||| : : : : : ||||| : : : : :
2854 TTTATC-----CGTGTAGCTACGGG 2874
Db ||||| : : : : : ||||| : : : : :
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768
Qy ||||| : : : : : ||||| : : : : :
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCAAGTGTCTGG 2934
Db ||||| : : : : : ||||| : : : : :
769 HisSerAspAsnArgMetGluThrHisTyThrSerLeuProGlnSerGluGlySerTrp 788
Qy ||||| : : : : : ||||| : : : : :
2935 GATAATACTGTCTGGTGGAGATGGAGCGGATTAACCGATGTGTGATTACTCTCATCT 2994
Db ||||| : : : : : ||||| : : : : :
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808
Qy ||||| : : : : : ||||| : : : : :
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGCAGCTGAGTTTCTTATCGGATCAT 3054
Db ||||| : : : : : ||||| : : : : :
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrglyValSerGln 828
Qy ||||| : : : : : ||||| : : : : :
3055 GAATCTTTTACAGAGAAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTTAAAT 3114
Db ||||| : : : : : ||||| : : : : :
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848
Qy ||||| : : : : : ||||| : : : : :
3115 CTATCAGTTCTGTGGAGTGAAGTTTGATCGAGTTGTAGTACATCATCTTAATAATAT 3174
Db ||||| : : : : : ||||| : : : : :
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyThrTrp 868
Qy ||||| : : : : : ||||| : : : : :
3175 AGCTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTAGAGCAACG 3234
Db ||||| : : : : : ||||| : : : : :
869 AspLeuSerGlyPhePheValSerAspValTyArgAsnAsnProGlnSerThrAlaThr 888
Qy ||||| : : : : : ||||| : : : : :
3235 CTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATAGCAACATGAGATT 3294
Db ||||| : : : : : ||||| : : : : :
535 IleGluGlyAsnIleTyrgluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553
Qy ||||| : : : : : ||||| : : : : :
2302 CTGAAATATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTGTACTTAGGG 2361
Db ||||| : : : : : ||||| : : : : :
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569
Qy ||||| : : : : : ||||| : : : : :
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397
Db ||||| : : : : : ||||| : : : : :
570 SerLeuIleProValProAlaAspProAsnSerGluTyrglyPheGlnGlnTrp 589
Qy ||||| : : : : : ||||| : : : : :
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457
Db ||||| : : : : : ||||| : : : : :
590 AsnValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrp 608
Qy ||||| : : : : : ||||| : : : : :
2458 ACTAAAACTGGTATAATCTCGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517
Db ||||| : : : : : ||||| : : : : :
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628
Qy ||||| : : : : : ||||| : : : : :
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577
Db ||||| : : : : : ||||| : : : : :
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648
Qy ||||| : : : : : ||||| : : : : :
2578 CGCTCTTATGTCGAGGATATGCGTTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637
Db ||||| : : : : : ||||| : : : : :
649 MetGluHisLysGlnGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 668
Qy ||||| : : : : : ||||| : : : : :
2638 GATGCTTATGTCGAGGATATCGGTATATAGTGGGGTTATCTCTAGGAGCAAACTCC 2697
Db ||||| : : : : : ||||| : : : : :
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyTyrglyValIleGlySerAla 688
Qy ||||| : : : : : ||||| : : : : :
2698 TACTTT---GGATCATCGATGTTTCTAGCATTTTACCGAAGTATTTGTTAGATCTAAA 2754
Db ||||| : : : : : ||||| : : : : :
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708
Qy ||||| : : : : : ||||| : : : : :
2755 GATTATGATGTCGTCTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814
Db ||||| : : : : : ||||| : : : : :
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrglyGlyThrLeuPhePheLysHis 728
Qy ||||| : : : : : ||||| : : : : :
2815 CAACAGCTTTA-----TGTGGATCTCTTTGTTGGAGATGG 2853
Db ||||| : : : : : ||||| : : : : :
729 SerHisThrLeuGlnProGlnAsnTyrglyLeuArgLeuGlyArgAlaLysPheSerGluSer 748
Qy ||||| : : : : : ||||| : : : : :
2854 TTTATC-----CGTGTAGCTACGGG 2874
Db ||||| : : : : : ||||| : : : : :
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768
Qy ||||| : : : : : ||||| : : : : :
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCAAGTGTCTGG 2934
Db ||||| : : : : : ||||| : : : : :
769 HisSerAspAsnArgMetGluThrHisTyThrSerLeuProGlnSerGluGlySerTrp 788
Qy ||||| : : : : : ||||| : : : : :
2935 GATAATACTGTCTGGTGGAGATGGAGCGGATTAACCGATGTGTGATTACTCTCATCT 2994
Db ||||| : : : : : ||||| : : : : :
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808
Qy ||||| : : : : : ||||| : : : : :
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGCAGCTGAGTTTCTTATCGGATCAT 3054
Db ||||| : : : : : ||||| : : : : :
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrglyValSerGln 828
Qy ||||| : : : : : ||||| : : : : :
3055 GAATCTTTTACAGAGAAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTTAAAT 3114
Db ||||| : : : : : ||||| : : : : :
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848
Qy ||||| : : : : : ||||| : : : : :
3115 CTATCAGTTCTGTGGAGTGAAGTTTGATCGAGTTGTAGTACATCATCTTAATAATAT 3174
Db ||||| : : : : : ||||| : : : : :
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyThrTrp 868
Qy ||||| : : : : : ||||| : : : : :
3175 AGCTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTAGAGCAACG 3234
Db ||||| : : : : : ||||| : : : : :
869 AspLeuSerGlyPhePheValSerAspValTyArgAsnAsnProGlnSerThrAlaThr 888
Qy ||||| : : : : : ||||| : : : : :
3235 CTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATAGCAACATGAGATT 3294
Db ||||| : : : : : ||||| : : : : :
535 IleGluGlyAsnIleTyrgluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553
Qy ||||| : : : : : ||||| : : : : :
2302 CTGAAATATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTGTACTTAGGG 2361
Db ||||| : : : : : ||||| : : : : :
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569
Qy ||||| : : : : : ||||| : : : : :
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397
Db ||||| : : : : : ||||| : : : : :
570 SerLeuIleProValProAlaAspProAsnSerGluTyrglyPheGlnGlnTrp 589
Qy ||||| : : : : : ||||| : : : : :
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457
Db ||||| : : : : : ||||| : : : : :
590 AsnValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrp 608
Qy ||||| : : : : : ||||| : : : : :
2458 ACTAAAACTGGTATAATCTCGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517
Db ||||| : : : : : ||||| : : : : :
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628
Qy ||||| : : : : : ||||| : : : : :
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577
Db ||||| : : : : : ||||| : : : : :
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648
Qy ||||| : : : : : ||||| : : : : :
2578 CGCTCTTATGTCGAGGATATGCGTTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637
Db ||||| : : : : : ||||| : : : : :
649 MetGluHisLysGlnGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 668
Qy ||||| : : : : : ||||| : : : : :
2638 GATGCTTATGTCGAGGATATCGGTATATAGTGGGGTTATCTCTAGGAGCAAACTCC 2697
Db ||||| : : : : : ||||| : : : : :
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyTyrglyValIleGlySerAla 688
Qy ||||| : : : : : ||||| : : : : :
2698 TACTTT---GGATCATCGATGTTTCTAGCATTTTACCGAAGTATTTGTTAGATCTAAA 2754
Db ||||| : : : : : ||||| : : : : :
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708
Qy ||||| : : : : : ||||| : : : : :
2755 GATTATGATGTCGTCTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814
Db ||||| : : : : : ||||| : : : : :
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrglyGlyThrLeuPhePheLysHis 728
Qy ||||| : : : : : ||||| : : : : :
2815 CAACAGCTTTA-----TGTGGATCTCTTTGTTGGAGATGG 2853
Db ||||| : : : : : ||||| : : : : :
729 SerHisThrLeuGlnProGlnAsnTyrglyLeuArgLeuGlyArgAlaLysPheSerGluSer 748
Qy ||||| : : : : : ||||| : : : : :
2854 TTTATC-----CGTGTAGCTACGGG 2874
Db ||||| : : : : : ||||| : : : : :
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768
Qy ||||| : : : : : ||||| : : : : :
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCAAGTGTCTGG 2934
Db ||||| : : : : : ||||| : : : : :
769 HisSerAspAsnArgMetGluThrHisTyThrSerLeuProGlnSerGluGlySerTrp 788
Qy ||||| : : : : : ||||| : : : : :
2935 GATAATACTGTCTGGTGGAGATGGAGCGGATTAACCGATGTGTGATTACTCTCATCT 2994
Db ||||| : : : : : ||||| : : : : :
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808
Qy ||||| : : : : : ||||| : : : : :
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGCAGCTGAGTTTCTTATCGGATCAT 3054
Db ||||| : : : : : ||||| : : : : :
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrglyValSerGln 828
Qy ||||| : : : : : ||||| : : : : :
3055 GAATCTTTTACAGAGAAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTTAAAT 3114
Db ||||| : : : : : ||||| : : : : :
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848
Qy ||||| : : : : : ||||| : : : : :
3115 CTATCAGTTCTGTGGAGTGAAGTTTGATCGAGTTGTAGTACATCATCTTAATAATAT 3174
Db ||||| : : : : : ||||| : : : : :
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyThrTrp 868
Qy ||||| : : : : : ||||| : : : : :
3175 AGCTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTAGAGCAACG 3234
Db ||||| : : : : : ||||| : : : : :
869 AspLeuSerGlyPhePheValSerAspValTyArgAsnAsnProGlnSerThrAlaThr 888
Qy ||||| : : : : : ||||| : : : : :
3235 CTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATAGCAACATGAGATT 3294
Db ||||| : : : : : ||||| : : : : :
535 IleGluGlyAsnIleTyrgluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553
Qy ||||| : : : : : ||||| : : : : :
2302 CTGAAATATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTGTACTTAGGG 2361
Db ||||| : : : : : ||||| : : : : :
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569
Qy ||||| : : : : : ||||| : : : : :
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397
Db ||||| : : : : : ||||| : : : : :
570 SerLeuIleProValProAlaAspProAsnSerGluTyrglyPheGlnGlnTrp 589
Qy ||||| : : : : : ||||| : : : : :
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457
Db ||||| : : : : : ||||| : : : : :
590 AsnValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrp 608
Qy ||||| : : : : : ||||| : : : : :
2458 ACTAAAACTGGTATAATCTCGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517
Db ||||| : : : : : ||||| : : : : :
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628
Qy ||||| : : : : : ||||| : : : : :
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577
Db ||||| : : : : : ||||| : : : : :
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648
Qy ||||| : : : : : ||||| : : : : :
2578 CGCTCTTATGTCGAGGATATGCGTTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637
Db ||||| : : : : : ||||| : : : : :
649 MetGluHisLysGlnGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 668
Qy ||||| : : : : : ||||| : : : : :
2638 GATGCTTATGTCGAGGATATCGGTATATAGTGGGGTTATCTCTAGGAGCAAACTCC 2697
Db ||||| : : : : : ||||| : : : : :
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyTyrglyValIleGlySerAla 688
Qy ||||| : : : : : ||||| : : : : :
2698 TACTTT---GGATCATCGATGTTTCTAGCATTTTACCGAAGTATTTGTTAGATCTAAA 2754
Db ||||| : : : : : ||||| : : : : :
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708
Qy ||||| : : : : : ||||| : : : : :
2755 GATTATGATGTCGTCTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814
Db ||||| : : : : : ||||| : : : : :
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrglyGlyThrLeuPhePheLysHis 728
Qy ||||| : : : : : ||||| : : : : :
2815 CAACAGCTTTA-----TGTGGATCTCTTTGTTGGAGATGG 2853
Db ||||| : : : : : ||||| : : : : :
729 SerHisThrLeuGlnProGlnAsnTyrglyLeuArgLeuGlyArgAlaLysPheSerGluSer 748
Qy ||||| : : : : : ||||| : : : : :
2854 TTTATC-----CGTGTAGCTACGGG 2874
Db ||||| : : : : : ||||| : : : : :
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768
Qy ||||| : : : : : ||||| : : : : :
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCAAGTGTCTGG 2934
Db ||||| : : : : : ||||| : : : : :
769 HisSerAspAsnArgMetGluThrHisTyThrSerLeuProGlnSerGluGlySerTrp 788
Qy ||||| : : : : : ||||| : : : : :
2935 GATAATACTGTCTGGTGGAGATGGAGCGGATTAACCGATGTGTGATTACTCTCATCT 2994
Db ||||| : : : : : ||||| : : : : :
789 SerAsnGluCysIleAlaGlyIleGlyLeuAspLeuPheValLeuSerAsnPro 808
Qy ||||| : : : : : ||||| : : : : :
2995 AAGCTTATTGATGATGCTGCTCTTCTCGTGCAGCTGAGTTTCTTATCGGATCAT 3054
Db ||||| : : : : : ||||| : : : : :
809 HisProLeuPheLysThrPheIleProGlnMetLysValGluMetValTyrglyValSerGln 828
Qy ||||| : : : : : ||||| : : : : :
3055 GAATCTTTTACAGAGAAAGCGCATCAAGCTCGGGCATTTCAAGAGCGGACATCTCTTAAAT 3114
Db ||||| : : : : : ||||| : : : : :
829 AsnSerPhePheGluSerSerSerAspGlyArgGlyPheSerIleGlyArgLeuLeuAsn 848
Qy ||||| : : : : : ||||| : : : : :
3115 CTATCAGTTCTGTGGAGTGAAGTTTGATCGAGTTGTAGTACATCATCTTAATAATAT 3174
Db ||||| : : : : : ||||| : : : : :
849 LeuSerIleProValGlyAlaLysPheValGlnGlyAspIleGlyAspSerTyThrTrp 868
Qy ||||| : : : : : ||||| : : : : :
3175 AGCTTATGCGGCTTATATCTGTGATGCTTATCGCACCATCTCTGCTGCTAGAGCAACG 3234
Db ||||| : : : : : ||||| : : : : :
869 AspLeuSerGlyPhePheValSerAspValTyArgAsnAsnProGlnSerThrAlaThr 888
Qy ||||| : : : : : ||||| : : : : :
3235 CTCTATCCATCAACAGACATGACAAACAGATGCTCTTCTTATAGCAACATGAGATT 3294
Db ||||| : : : : : ||||| : : : : :
535 IleGluGlyAsnIleTyrgluSerHis---MetPheSerHisAspGlnLeuPheSerLeu 553
Qy ||||| : : : : : ||||| : : : : :
2302 CTGAAATATACAGTTAGGAGCTAAGCCCCAGCTAATGCCCATCAGATTGTACTTAGGG 2361
Db ||||| : : : : : ||||| : : : : :
554 LeuLysIleThrVal-----AspAlaAspValAspThrAsnValAspIleSer 569
Qy ||||| : : : : : ||||| : : : : :
2362 AATGAGATGCTCT-----AAGTATGGCTATCAAGAAAGCTGG 2397
Db ||||| : : : : : ||||| : : : : :
570 SerLeuIleProValProAlaAspProAsnSerGluTyrglyPheGlnGlnTrp 589
Qy ||||| : : : : : ||||| : : : : :
2398 AAGCTTGGTGGGATCTTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGG 2457
Db ||||| : : : : : ||||| : : : : :
590 AsnValAsnTrpThrAspThrAlaThrAsn---ThrLysGluAlaThrAlaThrTrp 608
Qy ||||| : : : : : ||||| : : : : :
2458 ACTAAAACTGGTATAATCTCGGCTCAGCGAGTAGCTTCTTCTGTTCCAAATAGTTTA 2517
Db ||||| : : : : : ||||| : : : : :
609 ThrLysThrGlyPheValProSerProGluArgLysSerAlaLeuValCysAsnThrLeu 628
Qy ||||| : : : : : ||||| : : : : :
2518 TGGGATCCATTTTATAGATATACGATCTCGCATTCAGCAATTCAGCAAGTGTGGATGG 2577
Db ||||| : : : : : ||||| : : : : :
629 TrpGlyValPheThrAspIleArgSerLeuGlnLeuValGluIleGlyAlaThrGly 648
Qy ||||| : : : : : ||||| : : : : :
2578 CGCTCTTATGTCGAGGATATGCGTTTCTGGAGTTTCAATTTCTTCTATCATGACCGC 2637
Db ||||| : : : : : ||||| : : : : :
649 MetGluHisLysGlnGlyPheTrpValSerSerMetThrAsnPheLeuHisLysThrGly 668
Qy ||||| : : : : : ||||| : : : : :
2638 GATGCTTATGTCGAGGATATCGGTATATAGTGGGGTTATCTCTAGGAGCAAACTCC 2697
Db ||||| : : : : : ||||| : : : : :
669 AspGluAsnArgLysGlyPheArgHisThrSerGlyTyrglyValIleGlySerAla 688
Qy ||||| : : : : : ||||| : : : : :
2698 TACTTT---GGATCATCGATGTTTCTAGCATTTTACCGAAGTATTTGTTAGATCTAAA 2754
Db ||||| : : : : : ||||| : : : : :
689 HisThrProLysAspLeuPheThrPheAlaPheCysHisLeuPheAlaArgAspLys 708
Qy ||||| : : : : : ||||| : : : : :
2755 GATTATGATGTCGTCTCCATCATCATGCTTCATAGGATCCGTTTATCTATCTACC 2814
Db ||||| : : : : : ||||| : : : : :
709 AspCysPheIleAlaHisAsnAsnSerArgThrTyrglyGlyThrLeuPhePheLysHis 728
Qy ||||| : : : : : ||||| : : : : :
2815 CAACAGCTTTA-----TGTGGATCTCTTTGTTGGAGATGG 2853
Db ||||| : : : : : ||||| : : : : :
729 SerHisThrLeuGlnProGlnAsnTyrglyLeuArgLeuGlyArgAlaLysPheSerGluSer 748
Qy ||||| : : : : : ||||| : : : : :
2854 TTTATC-----CGTGTAGCTACGGG 2874
Db ||||| : : : : : ||||| : : : : :
749 AlalleGluLysPheProArgGluIleProLeuAlaLeuAspValGlnValSerPheSer 768
Qy ||||| : : : : : ||||| : : : : :
2875 TTTGGAATCAGCATATGAAAACCTCATATACATTTGCGAGAGAGCAAGTGTCTGG 2934
Db ||||| : : : : : ||||| : : : : :


```
Db      889  LeuValMetSerProAepSerTrpLysIleArgGlyAenLeuSerArgGlnAlaPhe 908
Qy      3295 GTGCTTATGAGGATCTATGATCTCTCTAACAAGTAATATAGAGTATATGGCCATGGA 3354
Db      909  LeuLeuArgGlySerAenAenTyrValTyrAenSerAenCysGluLeuPheGlyHisTyr 928
Qy      3355 AGATATAGTATCGAGATCGCTCTCGAGGCTATGTTGATGTCGAGGAGTAGAGTCGG 3414
Db      929  AlaMetGluLeuArgGlySerSerArgAenTyrAenValAepValGlyThrLysLeuArg 948
Qy      3415 TTC 3417
Db      949  Phe 949

RESULT 30
US-09-198-452A-470
; Sequence 470, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; NUMBER OF SEQ ID NOS: 1998-11-24
; SEQ ID NO 470
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-470

Alignment Scores:
Pred. No.: 2,73e-91 Length: 930
Score: 1083.00 Matches: 304
Percent Similarity: 43.3% Conservative: 160
Best Local Similarity: 28.4% Mismatches: 406
Query Match: 13.7% Indels: 202
DB: 2 Gaps: 29

US-10-701-844-1 (1-4435) x US-09-198-452A-470 (1-930)
Qy      382  ATGCAAGCTCTTCCATAAGTCTCTTCTTCAATGATCTAGCTTATCTTGCTGCTCT 441
Db      1  MetLysIleProLeuHisLysLeuLeuIleSerSerThrLeuVal---ThrProIleLeu 19
Qy      442  TTAATGGGGGGATATGCACGAAATC---ATGGTTCTCCTCAGGAAATTCAGATGG 498
Db      20  LeuSerIleAlaThrTyrGlyAlaAepAlaSerLeuSerProThrAepSerPheAepGly 39
Qy      499  GAGAGCTTAAGTATCATCTTCCCTACTGTTATAGGAGATCCGAGTGGGACTACTGTT 558
Db      40  AlaGlyLysSerThrPheThrProLysSerThr---AlaAepAlaAenGlyThrAenTyr 58
Qy      559  TTTCTCGAGGAGAGTAAACATTAATAAAATCTTGACAAATTCATTGCGAGCTTTGCTTTA 618
Db      59  ValLeuSerGlyAenValTyrIleAenAepAlaGlyLysGly---ThrAlaLeuThrGly 77
Qy      619  AGTTGTTTGGAACTTATATAGGAGTTTTACTGTTTTAGGAGGAGGACACTCGTTGACT 678
Db      78  CysCysPheThrGluThrThrGlyAepLeuThrPheThrGlyLysGlyTyrSerPheSer 97
Qy      679  TTCGAGACATACCGACTTCTCAAAATGGGCGAGCTCAAGTAATAGCGCTCGTATGGA 738
Db      98  PheAenThrValAepAlaGlySerAenAlaGlyAlaAalaserThrThrAlaAepLys 117
Qy      739  CTGTTTACTATTGAGGGTTTTAAAGAAATTATCTTTTCCAAATTCGAATTCATTCTGCC 798
Db      118  AlaLeuThrPheThrGlyPheSerAenLeuSerPhe-----11e 130
Qy      799  GTACTGCCCTGCGCAACGACTAATAAGGTAGCCAGACTCCGACGACACA-----849
```

```
Db      131  AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu 150
Qy      850  TCTACACCGTCTATGATGATCTATTTATCTTAAACAGATCTTTTGTCTACTCAATAATGAG 909
Db      151  AenLeuThrAepAenGlyThrIleLeuPheSerGlnAenVal-----164
Qy      910  AAGTTCATCTCTATAGTAATTTAGTCTCTGGAGATCGGGAGCTATAGATGCTAAGAGC 969
Db      165  -----SerAenGluAlaAenAenAenGlyGlyAlaIleThrAlaLysThr 179
Qy      970  TTAACGGTTCAGGAATTAGCAAGCTTTGTGCTTCCAAAGAAATATCTGCTCAAGCTGAT 1029
Db      180  LeuSerIleSerGlyAenThrSerSerIleThrPheThrSerAenSerAlaLysLysLeu 199
Qy      1030 GGGGAGCTTGTCAAGTAGTCACCAAGTTCTCTCTATGCTAAGCTAAGAGGCTCTTATGCC 1089
Db      200  GlyGlyAlaIleTyrSerSerAlaAlaSerIleSerGlyAenThrGlyGlnLeuVal 219
Qy      1090 TTTGTAGGAATGTTGCGGAGTAAGAGGGAGGAGGATGCTGCTGTCAGGATGGGCAG 1149
Db      220  PheMetAenAen---LysGlyGluThrGlyGlyAlaLeuGlyPheGlu-----235
Qy      1150 CAGGAGGTGTCATCATCTCTCTTCAACAGAGATCCAGTAGTAAGTTTTCAGAAATACT 1209
Db      236  -----AlaSerSerSerIleThrGlnAenSerSerLeuPhePheSerGlyAenThr 252
Qy      1210 GCGGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGGAGGATTTACTCTCAGGGAAC 1269
Db      253  AlaThrAep-----255
Qy      1270 GTTCTCTTCGAAATATGGAACCACTTTGTTCTCAACATGTTGCTTCTCTGTTTAC 1329
Db      255  -----255
Qy      1330 ATTGCTGCTAAGCAACCAAGTGGACAGGCTTCTTAACGAGTAATAATTACGAGAT 1389
Db      256  ---AlaAla-----GlyLys 259
Qy      1390 GGAGGAGCTATCTTCTGTAAGAATGGTGGCCAGCAGGATCCAAATACTCGGATCAGTT 1449
Db      260  GlyGlyAlaIleTyrCysGluLys-----ThrGlyGluThrProThrLeu 274
Qy      1450 TCCTTTGATGAGAGGAGTAGTTTCTTTAGTACGATGAGTCTGCTGGGAAGGGGA 1509
Db      275  ThrLeuSerGlyAenLysSerLeuThrPheAlaGluAenSerSerValThrGlnGly 294
Qy      1510 GCTATTTATGCCAAAAGCTCTCGGTTGCTAACTGTGGCCCTGTACAATTTTAAGGAAT 1569
Db      295  AlaIleCysAlaHisGlyLeuAepLysSerAlaAlaGlyProThrLeuPheSerAenAen 314
Qy      1570 -----ATCGCTAATGATGTTGGAGCGATTTATTTTAGGAGAACTCGAGAG 1614
Db      315  ArgCysGlyAenThrAlaAlaGlyLysGlyGlyAlaIleAlaIleAlaAepSerGlySer 334
Qy      1615 CTCAGTTATCTGCTGATATGAGATATTTATTTTCGATGGGAATCTTAAAGAAGACGCC 1674
Db      335  LeuSerLeuSerAlaAenGlnGlyAspIleThrPheLeuGlyAen-----349
Qy      1675 AAAGAGATGTCGCGGATTAATGGCGCTAACTGCTCTCAAGCCATTTTCATGGGA 1734
Db      350  -----ThrLeuThrSerThrSerAlaProThrSerThrArgAenAlaIleTyrLeuGly 367
Qy      1735 TCGGAGGAGGAAAAATACGACATTAAGAGCTAAAGCAGGCGCATCAGATCTCTTTAATGAT 1794
Db      368  SerSerAlaLysIleThrAenLeuArgAlaAlaGlnGlyGlnSerIleTyrPheTyrAep 387
Qy      1795 CCCATCGAGATGGCAACGGAAAT-----AACGAGCAGCGCAG 1833
Db      388  ProIleAlaSerAenThrThrGlyAlaSerAepValLeuThrIleAenGlnProAepSer 407
Qy      1834 TCTTCCAAACTTCTTAAAAATTAACGATGTTGAAGATACACAGGCGGATTTCTTTTGTCT 1893
Db      408  AenSerProLeu-----AspTyrSerGlyThrIleValPheSer 420
```


Query Match:	13.7%	Indels:	202
DB:	2	Gaps:	29
US-10-701-844-1 (1-4435) x US-09-438-185A-448 (1-938)			
QY	382	ATGCAACGCTCTTCCATAAGTCTCTTCAATGATTCAGCTTATCTTCTGCTCTCT	441
DB	9	MetIyIleProLeuHibLysLeuLeuIleSerSerThrLeuVal---ThrProIleLeu	27
QY	442	TAAATGGGGGGATGTCAGCAGAAATC---ATGGTTCCTCAAGGAATTTACGATGG	498
DB	28	LeuSerIleAlaThrTyrglyAlaAspAlaSerLeuSerProThrAspSerPheaspGly	47
QY	499	GAGAGTTAACTGATCATCTTCCCTACTGTTATAGAGATCCGAGTGGACTACTGTT	558
DB	48	AlaGlyGlySerThrPheThrProIlySerThr---AlaAspAlaAsnGlyThrAsnTy	66
QY	559	TTTTCTGAGAGAGTTAAACATTAATAAATCTTGACAATTCATTGACAGCTTTCCTTTA	618
DB	67	ValLeuSerGlyAsnValTyrlleAsnAspAlaGlyLysGly---ThrAlaLeuThrGly	85
QY	619	AGTTCTTTTGGGAACCTTATTAGGAGTCTTACTGTTTATAGGAGGAGCAGCTCGTTGACT	678
DB	86	CysCysPheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTySerPheSer	105
QY	679	TCGGAGAACATACGAGCTCTACAAATGGCGAGCTCTAAGTAATAGCGCTGCTGATGA	738
DB	106	PheAsnThrValAspAlaGlySerAsnAlaGlyAlaAlaAlaSerThrThrAlaAspLys	125
QY	739	CTGTTTACTATTGAGGGTTTAAAGAAATTAATCTCTTTTCCAAATTCGAATTCATTGCT	798
DB	126	AlaLeuThrPheThrGlyPheSerAsnLeuSerPhe	138
QY	799	GTAATGCTGCTGCAACGACTAATAAGGGTAGCCAGACTCCGACGACCAACA	849
DB	139	AlaAlaProGlyThrThrValAlaSerGlyLysSerThrLeuSerSerAlaGlyAlaLeu	158
QY	850	TCTACACCGTCTAATGGTACTATTTATCTTAAACAGATCTTTTGTACTCAATATGAG	909
DB	159	AsnLeuThrAspAsnGlyThrIleLeuPheSerGlnAsnVal	172
QY	910	AAGTCTCATCTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATCTAAGAGC	969
DB	173	-----SerAsnGluAlaAsnAsnGlyGlyAlaIleThrThrLysThr	187
QY	970	TTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAAAGAAATVACTGCTCAAGCTGAT	1029
DB	188	LeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLysLysLeu	207
QY	1030	GGGGAGCTGTGCAAGTAGTCACCAAGTTTCTGCTATGGCTAAGCAGGCTCTATTGCC	1089
DB	208	GlyGlyAlaIleTySerSerAlaAlaAlaSerIleSerGlyAsnThrGlyGlnLeuVal	227
QY	1090	TTTGTAGCGAAGTTGCGAGCTAAGAGGGGAGGATTTGCTGCTTCAGGATGGCGAG	1149
DB	228	PheMetAsnAsn---LysGlyGluThrGlyGlyGlyAlaLeuGlyPheGlu	243
QY	1150	CAGGAGGTGTCATCTACTCTCAACAGAAAGATCCAGTAGTAAGTTTTCAGAAATACT	1209
DB	244	-----AlaSerSerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThr	260
QY	1210	GGGTAGAGTTTGTAGGGAACTAGCCGAGTAGGAGGAGGATTTACTCTCAGCGAAC	1269
DB	261	AlaThrAsp	263
QY	1270	GTTGCTTCTCTGAATAATGGAATAACCTTGTTTCTCAACAATGTTGCTCTCTGTTTAC	1329
DB	263		263
QY	1330	ATTGCTGTAGCAACCAACAAAGTGGACAGGCTTCTTAATACGAGTAATAATTACGAGAT	1389
DB	264	---AlaAla	267
QY	1390	GGAGGAGCTATCTTCTGTAAGAATGGTGGCAAGAGATCCATAAATCTCGGATCACTT	1449
DB	268	GlyGlyAlaIleTyCysGluLys-----ThrGlyGluThrProThrLeu	282
QY	1450	TCCTTTGATGAGAGGAGTAGTTTCTTTAGTAGCAATGTAGTCTGGGAAAGGGGA	1509
DB	283	ThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThrGlnGlyGly	302
QY	1510	GCTATTTATGCAAAAAGCTCTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1569
DB	303	AlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPheSerAsnAsn	322
QY	1570	-----ATCGCTAATGATGGTGGAGGATTTATTTAGGAGAACTCGGAGAG	1614
DB	323	ArgCysGlyAsnThrAlaAlaGlyGlyGlyAlaIleAlaIleAlaAspSerGlySer	342
QY	1615	CTCAGTTTATCTGCTGATTTATGAGATATTTTTCGATGGGAATCTTTAAAGAACACCC	1674
DB	343	LeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsn-----	357
QY	1675	AAAGAGATGCTGCGGATGTTAATGGGTAATCTGCTGCTCACAAGCCATTCGATGGGA	1734
DB	358	-----ThrLeuThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyLeuGly	375
QY	1735	TCGGGAGGAAATAACGACATTAAGAGCTAAAGCAGGATCAGATTTCTTTAATGAT	1794
DB	376	SerSerAlaLysIleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyPheTyAsp	395
QY	1795	CCCATCGAGATGGCAACCGGAAAT-----AACCAACCCAGCGCAG	1833
DB	396	ProIleAlaSerAsnThrThrGlyAlaSerAspValLeuThrIleAsnGlnProAspSer	415
QY	1834	TCTTCCAACTCTTAAATAATTAACGATGGTGAAGATACACAGGGATTTGTTTCTGCT	1893
DB	416	AsnSerProLeu-----AspTySerGlyThrIleValPheSer	428
QY	1894	-----AATGGAAGCAGTACTTTGTAC	1914
DB	429	GlyGluLysLeuSerAlaAspGluAlaLysAlaAlaAspAsnPheThrSerIleLeuLys	448
QY	1915	CAAAATGTTACGATAGCAAGGAGGATTTCTTCTGTAAGGAGGCAAAATTTATCAGTG	1974
DB	449	GlnProLeuAlaLeuAlaSerGlyThrLeuAlaLeuLysGlyAsnValGluLeuAspVal	468
QY	1975	AATCTCTAAGTCAGACAGGCTGCTGATGATGGAAGCTGGGAGTACATGGGATTTT	2034
DB	469	AsnGlyPheThrGlnThrGluGlySerThrLeuLeu-----	480
QY	2035	GTAACCTCCACCAACCAACAGCTCTCCGCTAATCATGATTCAGCTCAGCTTTCCAAT	2094
DB	481	-----MetGlnProGlyThrLysLysLysAlaAspThrGluAlaIleSerLeuThrLys	498
QY	2095	CTGCATTTGCTCTCTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	2154
DB	499	LeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSer-----	513
QY	2155	CCTCAGCGCAAGATCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	2205
DB	514	-----IleGluThrAlaGlyAlaAsnLys	521
QY	2206	TCTGTTTACAAATTTAGTGGGCTATCTTTTGGAGGATTTGGATGATACAGCTTATGATGG	2265
DB	522	ThrIleThrLeuThrSerProLeuValPheGlnAspSerSerGlyAsnPheTyrglu---	540
QY	2266	TATGATGGCTAGGTTCTTAATCAAAAAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT	2325
DB	541	-----SerHisThrIleAsn-----GlnAlaPheThrGln	550
QY	2326	CCC-----CCAGCTAATGCCCCCATCAGATTTGACTCTTAGGG-----	2361
DB	551	ProLeuValPheThrAlaAlaThrAlaAlaSerAspIleTyIleAspAlaLeuLeu	570
QY	2362	-----AATGAGATGCTTAAGTATGCTATCAAGGAAGCTGGAAGCTTGGC	2406

US-09-198-452A-472
 ; Sequence 472, Application US/09198452A
 ; Patent No. 6559294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 472
 ; LENGTH: 927
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: 1...927
 ; OTHER INFORMATION: Xaa=unknown or other
 US-09-198-452A-472

Alignment Scores:
 Pred. No.: 4,48e-83 Length: 927
 Score: 994.50 Matches: 308
 Percent Similarity: 43.8% Conservative: 157
 Best Local Similarity: 29.0% Mismatches: 413
 Query Match: 12.6% Indels: 183
 DB: 2 Gaps: 38

US-10-701-844-1 (1-4435) x US-09-198-452A-472 (1-927)

QY	382	ATGCAACGCTTTCCATAGTCTTTCTTTCAATGATCTAGCTTATCTTGCTGCTCT	441
DB	1	MetLysSerSerLeuHisrPheLeuIleSerSerSerLeuAlaLeuPro---LeuSer	19
QY	442	TTAAATGGGGGGGATATGACAGCAAAATCATGGTT-----CCTCAAGGAATTTAC	492
DB	20	LeuAsnPheSerAlaPheAlaAlaValValGluIleAsnLeuGlyProThrAsnSerPhe	39
QY	493	GATGGGAG---ACGTTAACTGTATCATTTCCCTATCTATCTATAGGAGTCGAGTGGG	549
DB	40	SerGlyProGlyThrTyThr-----ProProAlaGlnThrThrAsnAlaAspGly	56
QY	550	ACTACTGTTTTTCTGCAGAGAGTTAAACATTAATAAAATCTTGACAAATCTATTGACGCT	609
DB	57	ThrIleTyAsnLeuThrGlyAspValSerIleThrAsnAla---GlySerProThrAla	75
QY	610	TTGCCCTTTAAGTTGTTTGGGAACCTTATTAGGAGTTTACTGTTTTAGGAGAGACAC	669
DB	76	LeuThrAlaSerCysPheLysGluThrThrGlyAsnLeuSerPheGlnGlyHisGlyTyr	95
QY	670	TCGTTGACTTTCAGAACATACGACTTCTACAAATGGGCGAGCTCTAAGTAATAGCGCT	729
DB	96	GlnPheLeuLeuGlnAsnIleAspAlaGlyAlaAsn---CysThrPheThrAsnThrAla	114
QY	730	GCTGATGAGCTGTTTACTATTAGGGGTTTAAAGAAATTAATCTTTTCCAAATTCGAATCA	789
DB	115	AlaAsnLysLeuLeuSerPheSerGlyPheSerTyLeuSerLeu-----	129
QY	790	TTACTTCCCGTACTGCTGCTGCAACGACTAATAAGGTAGCCAGACTCCGACGACACACA	849
DB	130	-----IleGlnThrThrAsn	134
QY	850	TCTACACCGTCTAATGCTACTATTATTCTAAACAGATCTTTTGTGTACTCAATAATGAG	909
DB	135	AlaThrThrGlyThrGlyAlaIleLysSerThrGlyAlaCysSerIleGlnSerAsnTy	154
QY	910	AAGTTCTCATTTATAGTAATTTAGTCTCTGAGATGGGGAGCTATAGATGCTAAGACG	969
DB	155	SerCysTyPheGlyGlnAsnPheSerAsnAspAsnGlyGlyAlaLeuGlnGlySerSer	174
QY	970	TTAACGGTTCAAGGAATTAGCAAGCTTTGTGTCTTCCAGAAATACTGCTCAAGCTGAT	1029

DB	175	IleSerLeu---SerLeuAsnProAsnLeuThrPheAlaLysAsnLysAlaThrGlnLys	193
QY	1030	GGGGAGCTTGTCAAGTAGTCACCAAGTTCTCTGCTATGGCTAACGAGGCTCTATTGCC	1089
DB	194	GlyGlyAlaLeu-----TyrSerThr	200
QY	1090	TTGTAGCGAATGTTGCAGGAGTAGAGGGGGAGGATTGCTGTGTTTCAGGATGGCAG	1149
DB	201	-----GlyGlyIleThr-----	204
QY	1150	CAGGAGTGTCAATCATCTACTTCAACAGAAGATCCAGTAGTAAGTTTTCAGAAATACT	1209
DB	205	-----IleAsnAsnThrLeuAsnSer-----AlaSerPheSerGluAsnThr	218
QY	1210	GGGTAGAGTTTGTAGGGAACGTAGCCGAGTAGGAGGAGGATTACTCTCTACGGGAAC	1269
DB	219	Ala-----AlaAsnAsnGlyGlyAlaIleTyThrGluAlaSer	231
QY	1270	GTTGCTTCTCGAATAATGAAACCTTG---TTTCTCAACAATGTTGCTTCTCTCTGTT	1326
DB	232	---SerPheIleSerSerAsnLysAlaIleSerPheIleAsnAsn-----	245
QY	1327	TACATTGCTCTAAGCAACCAACAGTGGACAGGCTTCTTAATACGAGTAATAATTACGGA	1386
DB	246	SerValThrAlaThrSerAlaThr-----	253
QY	1387	GATGAGAGCTATCTTCTGTGAAGATGCTGGCGCAAGCAGGATCCAAATACTCTGGATCA	1446
DB	254	---GlyGlyAlaIleTyCys-----SerSerThrSerAlaProLysProVal	268
QY	1447	GTTTCTTGTAGGAGGAGTAGTGTCTTTCTTAGTACCAATGTAGCTCTGGGAAAGG	1506
DB	269	LeuThrLeuSerAspAsnGlyGluLeuAsnPheIleGlyAsnThrAlaIleThrSerGly	288
QY	1507	GGAGCTATTATGCCAAAAGCTCTCGGTTGTCTAACTGTGGCCCTGTACAAATTTTAAAG	1566
DB	289	GlyAlaIleTyThrAspAsnLeuValLeuSerSerGlyGlyProThrLeuPheLysAsn	308
QY	1567	ATATCGCTAATGAT-----GCTGGAGCGATTTATTTAGGAGAATCTGGA	1611
DB	309	AsnSerAlaIleAspThrAlaAlaProLeuGlyGlyAlaIleAlaIleAlaSerGly	328
QY	1612	GAGCTCAGTTTATCTGCTATTATGAGATATTTTCGATGGGAATCTTAAAAAGCA	1671
DB	329	SerLeuSerLeuSerAlaLeuGlyGlyAspIleThrPheGluGlyAsn-----Thr	345
QY	1672	GCCAAAGAGAATGCTCCCGATGTTAATGGCGTAATGCTCTCTCAACGCCATTCGATG	1731
DB	346	ValValLysGlyAlaSer-----SerSerGlnThrThrThrArgAsnSerIleAsnIle	363
QY	1732	GGA---TCGGGAGGGGAAATAACGACATTAAAGAGCTAAAGCAGGCGATCAGATCTCTT	1788
DB	364	GlyAsnThrAsnAlaLysIleValGlnLeuArgAlaSerGlnGlyAsnThrIleTyPhe	383
QY	1789	AATGATCCCATCGAGTGCACAAAGCAATAACACCGCAGCGAGCTTCCAAACTCTA	1848
DB	384	TyrAspProIleThr-----ThrSerIleThrAlaAlaLeuSerAspAlaLeu	399
QY	1849	AAAAATTAACCATGGTGAA-----GGATACACAGGGGATATTGTTTGTCT	1893
DB	400	AsnLeuAsnGlyProAspLeuAlaGlyAsnProAlaTyGlnGlyThrIleValPheSer	419
QY	1894	-----AATGGAAGCAGTACTTGTGATC	1914
DB	420	GlyGluLysLeuSerGluAlaGluAlaAlaGluAlaAspAsnLeuLysSerThrIleGln	439
QY	1915	CAAAATGTTACGATAGCAAGGAGGATTGTTCTTCGTGAAAAGGCAAAATATTCAGTG	1974
DB	440	GlnProLeuThrLeuAlaGlyGlnLeuSerLeuLysSerGlyValThrLeuValAla	459
QY	1975	AATTTCTTAAGTCAGACAGGTGGAGT---CTGTATATGGAAGCTGGGAGTACATGGAT	2031

326 SerPheThrSerAsnThrThrAlaAlaAsnGlyGlyAlaIleTyrAlaThrLysCysThr 345
Qy
475 GTTCTCAAGGAATTACGATGGGAGAGCTTAAGTATGATATCATTTCCCTACTACTGTTATA 534
Db
346 Leu-----AspGlyAsnThr-----ThrLeuThrPheAspGlnAsnThrAla 359
Qy
535 GGAGATCCGAGTGGGACTACTGTTTCTGCGAGGAGAGTTAAACATTAAAAAATCTTCAC 594
Db
360 ThrAlaGlyCysGlyAlaIleTyrThrGluThrGlu----- 372
Qy
595 AATTCTATTGCAGCTTTCCTTTAAGTTGTTTGGGAACCTTATTAGGAGATTTCAGTT 654
Qy
373 -----AspPheSerLeu 376
Qy
655 TTAGGAGAGACACTCGTTGACTTTCGAGAACATACGGACTTCACAAATGGGCGAGCT 714
Db
377 LysGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThrGlyGlyAlaLeu 396
Qy
715 CTAAGT-----AATAGCGCTGCTGATGGA-----CTGTTTACTATTGAGGT 756
Db
397 TyrSerLysGlyAsnSerSerLeuThrGlyAsnThrAsnLeuLeuPheSer-----Gly 414
Qy
757 TTTAAGAATTATCCCTTTTCCAAATTCATTACTTTCGCTACTGCTGCTGCAACG 816
Qy
415 AsnLysAlaThrGlyProSerAsn-----SerSer 424
Qy
817 ACTAATAAGGTAGCCAGACTCCGAGCACACATCTACACCTCTAATGGTACTATTAT 876
Db
425 AlaAsnGlnGluGlyCysGlyAlaIleLeuAlaPheIleAspSerGlySerValSer 444
Qy
877 TCTAAAAACAGATCTTTTGTACTCAATATGAGAATCTCTCAATCTATAGTAATTTAGTC 936
Db
445 AspLysThrGlyLeuSerIleAlaAsnAsnGlnGluValSerLeuThrSerAsnAlaAla 464
Qy
937 TCTGAGATGGGGAGCTATAGATCTTAAGACTTAACGGTTCAAGGAATTAGCAAGCTT 996
Db
465 ThrValSerGlyAlaIleTyrAlaThrLysCysThrLeuThrGlyAsnGlySerLeu 484
Qy
997 TGTGCTTCCAGAAATCTCTCAAGCTGATGGGGAGCTTGTCAGTAGTACC----- 1053
Db
485 ---ThrPheAspGlyAsnThrAlaGlyThrSerGlyGlyAlaIleTyrThrGluThrGlu 503
Qy
1054 AGTTTCTGCTATGCTTAACAGGCTCTCTATTGCTTGTAGCGAATGTCAGGAGTA 1113
Db
504 AspPheThrLeuThrGlySerThrGlyThrValThrPheSerThrAsnThrAlaLysThr 523
Qy
1114 AGAGGGGGAGGATT----- 1128
Db
524 ---GlyGlyAlaLeuTyrSerLysGlyAsnAsnSerLeuSerGlyAsnThrAsnLeuLeu 542
Qy
1129 -----GCTGCTGTTCCAGATGGCGAGCAG 1152
Db
543 PheSerGlyAsnLysAlaThrGlyProSerAsnSerSerAlaAsnGlnGluGlyCysGly 562
Qy
1153 CGAGTG-----TCATCATCTACTTCAACA----- 1176
Db
563 GlyAlaIleLeuSerPheLeuGluSerAlaSerValSerThrLysGlyLeuTrpIle 582
Qy
1177 ---GAAGATCCAGTAGTAGTTTTCAGAAATATCTGCG----- 1212
Db
583 GluAspAsnGluAsnValSerLeuSerGlyAsnThrAlaThrValSerGlyGlyAlaIle 602
Qy
1213 -----GTAGATTTCATGGAGACGTAGCC 1236
Db
603 TyrAlaThrLysCysAlaLeuHisGlyAsnThrThrLeuThrPheAspGlyAsnThrAla 622
Qy
1237 CGA---GTAGGAGGGGATTAC-----TCCTAC 1263
Db
623 GluThrAlaGlyGlyAlaIleTyrThrGluThrGluAspPheThrLeuThrGlySerThr 642
Qy
1264 GGGAGCTTGTCTTCTGTAATAT-----GGA 1290
Db
643 GlyThrValThrPheSerThrAsnThrAlaLysThrAlaGlyAlaLeuHisThrLysGly 662

1291 AAAACCTTGTTTCTCAACAATGTTGCTTCTCTCTGTTTACATTGCTGTAAGCAACCAACA 1350
Qy
663 AsnThrSerPheThrLysAsnLysAla-----LeuValPheSerGlyAsnSerAla 679
Db
1351 AGTGACAGGCTTCTTAATACGAGTAATAATTAACGAGATGAGAGAGCTATCTCTCTGT--- 1407
Qy
680 ThrAlaThrAlaThrThrThrAspGlnGluGlyCysGlyGlyAlaIleLeuCysAsn 699
Db
1408 -----AAGAATGGTCGCAAGCA 1425
Qy
700 IleSerGluSerAspIleAlaThrLysSerLeuThrLeuThrGluAsnGluSerLeuSer 719
Db
1426 GGATCCCAATAACTCTCGCATCAGTTTTC-----TTT 1455
Qy
720 PheIleAsnAsnThrAlaLysArgSerGlyGlyGlyIleTyrAlaProLysCysValIle 739
Db
1456 GATGAGAGGAGTAGTATTTCTTTAGTAGCAATAGTAGCTGCGGAAAGGGGAGCTATT 1515
Qy
740 SerGlySerGluSerIleAsnPheAspGlyAsnThrAlaGluThrSerGlyGlyAlaIle 759
Db
1516 TATGCCAAAAGCTCTCGTTCCTAACCTGTGCGCTGTACAAATTTTAAAGGAATATCGCT 1575
Qy
760 TyrSerLysAsnLeuSerIleThrAlaAsnGlyProValSerPheThrAsnAsnSerGly 779
Db
1576 AATGATGCTGAGCGCATTTATTTAGGAGAACTCTGAGAGCTCAGTTTATCTCTGATTAT 1635
Qy
780 GlyLysGlyGlyAlaIleTyrIleAlaAspSerGlyGluLeuSerLeuGluAlaIleAsp 799
Db
1636 GGAGATATATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGAAATGCTCCGATGTT 1695
Qy
800 GlyAspIleThrPheSerGlyAsn-----ArgAlaThrGluGlyThrSerThrPro 816
Db
1696 AATGCGGTAACTGCTCTCACAGCCATTTCCATGGGATCGGAGGGAATAACGACA 1755
Qy
817 Asn-----SerIleHisLeuGlyAlaGlyAlaLysIleThrLys 829
Db
1756 TTAAGAGCTTAAGCAGGCGCATCAGATTCTCTTTAATATGATCCCATCGAGATG----- 1806
Qy
830 LeuAlaAlaAlaProGlyHisThrIleTyrPheTyrAspProIleThrMetGluAlaPro 849
Db
1807 GCAAAACGGA----- 1815
Qy
850 AlaSerGlyGlyThrIleGluGluLeuValIleAsnProValValLysAlaIleValPro 869
Db
1816 -----AATAACCGCCAGCGCAGCTCTCCAACTTCTCAAATTAACGATGCT 1863
Qy
870 ProProGlnProLysAsnGlyProIleAlaSerValProValProValAlaProAla 889
Db
1864 GAAGGATACACAGGGGATATCTTTT----- 1890
Qy
890 AsnProAsnThrGlyThrIleValPheSerSerGlyLysLeuProSerGlnAspAlaSer 909
Db
1891 -----GCTAATGGAAGCAGTACTTGTACCAAAATGTTACGATAGACAGCAAGAGATT 1944
Qy
910 IleProAlaAsnThrThrThrIleLeuAsnGlnLysIleAsnLeuAlaGlyGlyAsnVal 929
Db
1945 GTTCTTCTGTAAGCAAAATTAATCAGTGAATTTCTTAAGTCAGACAGGTGGAGCT--- 2001
Qy
930 ValLeuLysGluGlyAlaThrLeuGlnValTyrSerPheThrGlnGlnProAspSerThr 949
Db
2002 CTGTATATCGAGCTGGGAGTACATGGGATTTGTAACCTCCACACCAACCAACAGCCT 2061
Qy
950 ValPheMetAspAlaGlyThrThrLeuGluThrThrThrThrAsn----- 964
Db
2062 CTGCGCGCTAATCAGTTGATCAGCTTTCCAAATCTGCATTTGTCTCTTCTTTCTTTGTTA 2121
Qy
965 ---AsnThrAspGlySerIleAspLeuLysAsnLeuSerValAsnLeuAspAlaLeuAsp 983
Db
2122 GCAAACATGCGATTGAGTAATCTCTACCAATCTCCAGCGCAAGATTCTCATCTCTGCA 2181
Qy
984 GlyLysArgMetIleThr----- 989
Db


```
QY 1204 AATACTGCGTAGAGTTTGTATGCGAAGCTAGCCGAGTAGGAGGAGGATTTACTCTCTAC 1263
Db 77 -----GlyAlaAlaIlePheThrGlu 83
QY 1264 GCGAAGCTTGCTTCTGTAATAATGGAAGAACCTTGTCTCAACAATGCTTCTCTCTCT 1323
Db 84 AlaSerValThrIleSerAsnAlaIleValSerPheIleAspAsn-----99
QY 1324 GTTTACATTGCTGCTAAGCAACCAACAGTGGACAGGCTCTTAATACGAGTAATAATTAC 1383
Db 100 -----LysValThrGlyAlaSerSerThrThrGlyAspMet 112
QY 1384 CGAGATGGAGGAGCTATCTTCTGTAAGAAATGGTGGCAAGCAGGATCCAATAACTGGA 1443
Db 113 -----SerGlyGlyAlaIleCysAlaTyrLys-----ThrSerThrAspThr 126
QY 1444 TCAGTTTCTCTTGTATGAGGAGGAGTAGTATTTCTTTAGTAGCAATGAGTCTGCGGAAA 1503
Db 127 LysValThrLeuThrGlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrAla 146
QY 1504 GGGGAGCTATTATGCCAAAGCTCTCGGTGCTAACTGGCCCTGTACAAATTTTA 1563
Db 147 GlyGlyAlaIleTyrValLysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSer 166
QY 1564 AGCAATATCGTAAT-----GATGGTGGAGCGATTTATTATTAGGAGAACTCT 1608
Db 167 ArgAsnSerValAsnGlyThrAlaProLysGlyGlyAlaIleAlaIleAlaIleGluAspSer 186
QY 1609 GGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTTTCGATGGGAATCTTAAAGA 1668
Db 187 GlyGluLeuSerLeuSerAlaAspSerGlyAspIleValPheLeuGlyAsn-----203
QY 1669 ACAGCCAAAGAGAACTGCTGCGGATGTAATGGCTGAATCTGTGTCCTCACAGCCATTCG 1728
Db 204 ThrValThrSerThrThrProGlyThrAsn-----ArgSerSerIleAsp 218
QY 1729 ATGGGATCGGGGGAATAACGACATTAAGAGCTAAAGCAGGCGATCAGATTCCTTT 1788
Db 219 LeuGlyThrSerAlaLysMetThrAlaLeuArgSerAlaAlaGlyArgAlaIleTyrPhe 238
QY 1789 AATGATCCCATCCAGATGCGCAAGGAAATAACCGCCAGCGAGTCTTCCAAACTCTA 1848
Db 239 TyrAspProIleThrThrGlySerSerThr-----ValThrAspValLeu 254
QY 1849 AAAATTAACGATGGTGAAGGA-----TACAGGGGATATGTTTGTCT 1893
Db 255 LysValAsnGluThrProAlaAspSerAlaLeuGlnTyrThrGlyAsnIleIlePheThr 274
QY 1894 -----AATGNAAGCAGTACTTTGTAC 1914
Db 275 GlyGluLysLeuSerGluThrGluAlaAlaAspSerLysAsnLeuThrSerLysLeuLeu 294
QY 1915 CAAAATGTTACGATAGACGAAGGAGTGTCTTCGTGAAAAGGCAAAATTAATCAGTG 1974
Db 295 GlnProValThrLeuSerGlyGlyThrLeuSerLeuLysHisGlyValThrLeuGlnThr 314
QY 1975 AATTCTCTAAGTCAGACAGGTGGAGT---CTGTATATGAAGCTGGAGTACATGGAT 2031
Db 315 GluAlaPheThrGlnGlnAlaAspSerArgLeuGluMetAspValGlyThrThrLeuGlu 334
QY 2032 TTTGTAATCCACACACCACCAACACCCCTCTGCGCTAATCAGTTGATCAGCTTTC 2091
Db 335 -----ProAlaAspThrSer-----ThrIleAsn 342
QY 2092 AATCTGCATTTGCTCTTCTTCTTCTTGTAGCAAAACAATGCAAGTACCAATCTCTTACC 2151
Db 343 AsnLeuValIleAsnIleSerSerIle-----351
QY 2152 AATCTCCAGCGCAAGATTTCTACCTGCAATCTGCTAGTAC-----ACAAGTGGTGTCT 2208
Db 352 -----AspGlyAlaLysAlaLysIleGluThrLysAlaThrSerLysAsn 367
QY 2209 GTTACAATTAGTGGGCTATCTTTTGTGAGGATTTGGATGATACAGCTTATGAGTAT 2268
Db 367 -----367
QY 368 LeuThrLeuSerGlyThrIleThrLeuLeuAspProThrGlyThrPheTyrGluAsnHis 387
QY 2269 GATTGGCTAGGTTCTTAATCAAAAATCAATGTCCTGAAATTAACAGTTA---GGGACTAAG 2325
Db 388 Ser-----LeuArgAsnProGlnSerTyrAspIleLeuGluLeuLysAlaSerGlyThrVal 406
QY 2326 CCCCAGCTAATAGCCCATCAGATTTGACTCTAGGGAATGAGATGCCCTAAGTATGCTAT 2385
Db 407 ThrSerThrAlaValThrProAspProIleMetGlyGluLysPhe---HisTyrGlyTyr 425
QY 2386 CAAGGAAGCTGG---AAGCTTGGTGGATCTTAATACAGCAAAATAATGCTCTTATCTAT 2442
Db 426 GlnGlyThrTyrProIleValThrGlyThrGlyAlaSerThrThrAlaThrPhe---444
QY 2443 CTGAAAGCTACAGTACGACTGAACTGAGTATTAATCTGGGCCCTGAGGAGTAGTCTCTTTG 2502
Db 445 -----AsnTyrThrLysThrGlyTyrIleProAsnProGluArgIleGlySerLeu 461
QY 2503 GTTCCAAATAGTTTATGGGATCCATTTTATAGATATACGATCTGCGCATTCAGCAATTC 2562
Db 462 ValProAsnSerLeuTyrAsnAlaPheIleAspIleSerSerLeuHisTyrLeuMetGlu 481
QY 2563 GCAAGTGTGAGTGGGCGCTCTTATTCGAGGATTAATGGGTTTCTGAGTTTCGATTC 2622
Db 482 ThrAlaAsnGluGlyLeuGlnGlyAspArgAlaPheTyrCysAlaGlyLeuSerAsnPhe 501
QY 2623 TTTCTATCATGACCGCATCTTTAGTTCAGGATTCGCTATATTAATAGTGGGCTTATTC 2682
Db 502 PheHisLysAspSerThrLysThrArgArgGlyPheArgHisLeuSerGlyGlyTyrVal 521
QY 2683 TTAGGAGCAACTCTCTACTTTTGGATCA---TCGATGTTTGGTCTAGCATTTATCCGA 2739
Db 522 IleGlyGlyAsnLeuHisThrCysSerAspLysIleLeuSerAlaAlaPheCysGlnLeu 541
QY 2740 TTTGGTAGATCTAAAGATTAATGATGTGTCTTCAATCATCATCTGTCGATAGATTC 2799
Db 542 PheGlyArgAspArgAspTyrPheValAlaLysAsnGlnGlyThrValTyrGlyGlyThr 561
QY 2800 GTTATCTATCTACCCCAACAGCTTTATGCGATCTTATGCGATCTTATGCGAGTGGTTATC 2859
Db 562 LeuTyr-TyrGlnHisAsnGluThrTyrIleSerLeuProCysLysLeuArgProCysSe 581
QY 2860 CGTGCT-----AGCTACCGGTT 2876
Db 581 rLeuSerTyrValProThrGluIleProValLeuPheSerGlyAsnLeuSerTyrThrHi 601
QY 2877 TGGGAATCAGCATATGAAAACCTCATATACATTTGAGAGGAGGAGCATGTTCTGGTGG 2936
Db 601 sThrAspAsnAspLeuLysThrLysTyrThrTyrProThrValLysGlySerTrpG1 621
QY 2937 TAATAACTGTCTGCTGGAGAGATTTGAGCGGATTTACCGATTTGATGTTACTCCATCTAA 2996
Db 621 yAsnAspSerPheAlaLeuGluPheGlyGlyArgAlaProIleCysLeuAspGluSerAl 641
QY 2997 GCTCTATTTGAATGAGTTCGCTCTTTCGTGCAAGCTGAGTTCCTTATCCCATCATGA 3056
Db 641 aLeuPhe---GluGlnTyrMetProPheMetLysLeuGlnPheValTyrAlaHisGlnG1 660
QY 3057 ATCTTTTACAGAGAAAGCGCATCAAGCTCGGCGATTTCAAGAGCGGACATCTCTTAATCT 3116
Db 660 uGlyPheLysGluGlnGlyThrGluAlaArgGluPheGlySerSerArgLeuValAsnLe 680
QY 3117 ATCAGTTCTCTGTTGGAGTGAAGTTTGAATTCATGTTCTTAGTACATCTTAATAATATAG 3176
Db 680 uAlaLeuProIleGlyIleArgPheAspLysGluSerAspCysGlnAspAlaThrTyrAs 700
QY 3177 CTTTATGCGGCTTATCTGTATGCTTATCGACCATCTTATCGACCATCTCTGGTACTGAG 3236
Db 700 nLeuThrLeuGlyTyrThrValAspLeuValArgSerAsnProAspCysThrThrLe 720
QY 3237 CCTATCCCATCAAGAGACATGGCAACAGATGCTCTTTCATTTAGCAAGACATGGAGTGT 3296
Db 3296 -----3296
```



```
Db 720 wArgIleSerGlyAspSerTrpLysThrPheGlyThrAsnLeuAlaArgGlnAlaLeuVa 740
Qy 3297 GGTAGAGATCTATGATGCTTCTTAACAGTAATATAGAGATATATGAGCCATGGAAG 3356
Db 740 lLeuArgAlaGlyAenHisPheCysPheAsnSerAsnPheGluAlaPheSerGlnPheSe 760
Qy 3357 ATATGAGTATCGAGATGCTTCTCGAGGCTATGTTTGTAGTCAGGAAGTAGTCCGGTT 3416
Db 760 rPheGluLeuArgGlySerSerAgaSntYrAsnValAspLeuGlyAlaLysTrpGlnPh 780
Qy 3417 C 3417
Db 780 e 780

RESULT 37
US-09-438-185A-6
; Sequence 6, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CFP0005
US-09-438-185A-6

Alignment Scores:
Pred. No.: 4,456-81 Length: 932
Score: 973.00 Matches: 306
Percent Similarity: 41.5% Conservative: 148
Best Local Similarity: 28.0% Mismatches: 386
Query Match: 12.3% Indels: 253
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x US-09-438-185A-6 (1-932)
Qy 375 TTGTGTGATGCAAGCTCTTCCATAAGTTCTTCTTCAAT-----GAT 419
Db 15 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 31
Qy 420 TCTAGCTATTCTGCTGCTCTTTAAATGGGGGGGATATGAGCAGAAATCATGGTTC 479
Db 32 ThrSerLeu-SerAlaThrThrIleSer-----LeuThr-Pr 43
Qy 480 TCAAGGATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATCTGTTATAGGAGA 539
Db 43 oGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTrpAsnVal----- 60
Qy 540 TCCGAGTGGACTACTGTTTTCGACGAGAGTTAACTATAAAATCTTGACAATTC 599
Db 61 -GlnAlaGlyAspValTrpSerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 80
Qy 600 TATTCGAGCTTTGCTTTAAAGTTGTTTGGGAACCTATTATAGGAGTTTCTGTTTAGG 659
Db 80 r-----AlaLeuAsnLysAlaCysPheAsnValThrSerGlySerValThrPheAlaGl 98
Qy 660 GAGAGGACACTGTTGACTTTTCGAGAACATACGGACTTCTACA-----AATGGGGCAGC 713
```

```
Db 98 yAsnHisGlyLeuTrpPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 118
Qy 714 TCTAAGTAATAGCGCTGCTGATGACTGTTTACTATTTAGGGGTTTTAAAGAAATATCTTT 773
Db 118 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 138
Qy 774 TTCCAATTGCNAATTCATTACTTCCGCTACTGCTGCTGCAACGACTAATAAGGGTAGCCA 833
Db 138 e-----lIeGl 140
Qy 834 GACTCCGACGACAACATCTACACCGTCTAATGGTACTATTATTCTATAAACACAGATCTTT 893
Db 140 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTrpSerLysAsnAlaLeuWe 158
Qy 894 GTTACTCAATAATAGAGAGTTCTCATTTATAGTAATTTAGTCTCTGAGAGATGGGGAGC 953
Db 158 tLeuLeuAsnAsnTrpValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 178
Qy 954 TATAGATGCTAAGAGCTTAACGGTTCAAGGATTTAGCAAGCTTTGTCTCTCCAGAAAA 1013
Db 178 aIleSerGlyAlaAsnValThrIleValGlyAsnTrpAspSerValSerPheTrpGlnAs 198
Qy 1014 TACTGCTCAAGCTGATGGGGAGCTTGTCAAGTAGTCACCAAGTTTCTCTGCTATGGCTAA 1073
Db 198 nAlaAlaThrPheGlyGly-----AlaIleHisSe 208
Qy 1074 CGAGGCTCTATGCCCTTTGTAGGAATTTGTGAGGAGTAAAGAGGGGAGGATGCTGCG 1133
Db 208 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 222
Qy 1134 TGTTCAAGATGGGAGGAGGAGTGTCTATCTTCAACAGAGATCCAGTAGTAAG 1193
Db 222 ----- 222
Qy 1194 TTTTCCAGAAATACTGCGTAGAGTTTGTATGGAAACGTAGCCCGAGTAGGAGGAGGAT 1253
Db 223 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 236
Qy 1254 TACTCTCTACGGGAACGTTGCTTCTCTGAATTAATGAAACCTTTGTTTCTCAACAATGT 1313
Db 236 uTrp-SerAspGlyAspIleAspGlnAsnAlaTrpValLeuPheArgGluAsnGl 256
Qy 1314 TGCTTCTCTGTTTACATTTGCTGTCTAGCAACCAACAGTGGACAGGCTTCTAATACGAG 1373
Db 256 uAlaLeuThrThrAlaIle----- 262
Qy 1374 TAATAATTACCGAGATGGAGGAGCTATCTTCTGTAAAGAAATGTTGCGCAAGGAGTCCA 1433
Db 263 -----GlyLysGlyGlyAlaValCysCys-----LeuPr 272
Qy 1434 TAACCTCTGATCA-----GTTTCTTGTATGAGAGGAGTAGTTT 1475
Db 272 oThr-SerGlySerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 292
Qy 1476 CTTTAGTAGCAATGTAGCTGCTGGGAAGGGGAGCTATTATTCGCAAAAGCTCCGCT 1535
Db 292 lPheGluArgAsnHisSerIleMetGlyGlyGlyAlaIleTrpAlaArgLysLeuSerIl 312
Qy 1536 TGCTAACTGTGCGCTGTACAAATTTTAAAGAAATTCGCT-----AATGA 1580
Db 312 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTrpAlaAsnSerGlnAsnLe 332
Qy 1581 TGGTGGAGCGGATTTATTTAGGAGAAATCTGGAGAGCTCAGTTTATCTGCTGATATGAGA 1640
Db 332 uGlyGlyAlaIleAlaIleAspThrGlyGlyGlyLysSerLeuSerAlaGluLysGlyTh 352
Qy 1641 TATTATTTTCGATGGGAATCTTAAAGAACAGCCAAAGAGATGCTGCCGATGTTAATGG 1700
Db 352 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGl 367
Qy 1701 CGTAACTGTGCTCTCAACAAGCCATTTCCATGGGATCGGGAGGAGAAATAACGACATTAAG 1760
Db 1760
```

Db 367 ylleHisLeuLeuGlnAenAla-----LysPheLeuLysLeuG1 380
 QY 1761 AGCTAAAGCAGGCGATCAGATCTCTTTAATGATCCATCGAGATGGCAACGGAATAA 1820
 Db 380 nAlaArgAenGlyTySerIleGluPheTyArgProfile-----Th 394
 QY 1821 CCAGCCAGGCGAGTCTTCCAAACTTCTAAAAATTAAAC-----GATGGTGAAGGATA 1871
 Db 394 rSerGluAlaAspGlySerThrGlnLeuAenIleAenGlyAspProLysAenLysGluTy 414
 QY 1872 CACAGGGGATATGTTTTGCT-----AATGGAG 1901
 Db 414 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAenAspProArgAspPheTy 434
 QY 1902 CAGTACTTTGTACCAAAATGTTACGATAGCAGCAAGAGATGTTCTTCGTGAAAGGC 1961
 Db 434 sSerThrIleProGlnAenValAenLeuSerAlaGlyTyLeuValIleGlyGluGlyAl 454
 QY 1962 AAAATTATCAGTGAATCTCTAAGTCAGACAGTGGAGT---CTGTATATGGAAGCTGG 2018
 Db 454 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuG1 474
 QY 2019 GAGTACATGGGATTTGTAACTCCACACCACACAGCCCTCTGCGCTAATCAGTT 2078
 Db 474 yThr-----LysLeuIleAlaSerLysGluAs 483
 QY 2079 GATCAGCTTTCCAAATCTGCATTTGCTCTTTCTTCTTCTTTGTTAGCAAAATGCGATTAC 2138
 Db 483 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 496
 QY 2139 GAATCTCTCCAAATCTCCAGCGCAAGATTTCTCATCTGCGAGTATTGGTAGCACAA 2198
 Db 497 -----SerSerSerThrAlaAlaValIleLysAlaAenTh 509
 QY 2199 TGCTGCT-----TCTGTACAATTAGTGGGCTATCTTT----- 2232
 Db 509 rAlaAenLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAenAl 529
 QY 2233 -TTTGAGGATTTGGATGATACAGCTTATGATGATGATGATGATGATGATGATGATGAT 2291
 Db 529 aTyrGluAspLeuArg-----MetArgAenSerGlnTh 540
 QY 2292 AATCAATGCTCTGAATATACAGTTAGGACTAGCCCAAGCTTAAATGCCCATCAGATT 2351
 Db 540 rPheProLeuLeuSerLeuGlu-----ProGlyAlaGlyGlySerValThrVa 556
 QY 2352 GACTCTAGGGAATGAGATG-----CCTAAGTATGGCTATCAAGGAAGCTGGAAGCT 2402
 Db 556 lThrAlaGlyAspPheLeuProValSerProHisTyTyGlyPheGlnGlyAenTrpLysLe 576
 QY 2403 TGGGTGGGATCCTAATACAGCAAAATATGCTCTTATCTCTGAAAGCTACATGGACTAA 2462
 Db 576 uAlaTrpThrGlyThrGlyAenLysValGlyGluPhe-----PheTrpAspLy 592
 QY 2463 AACTGGGTATATCTCGGGCTGAGCGAGTAGCTCTTTGTTCCAAATAGTTATGGGG 2522
 Db 592 sIleAenTyLysProArgProGluLysGluGlyAenLeuValProAenIleLeuTrpG1 612
 QY 2523 ATCCATTTTATAGATATAGCTGGCATTTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 2582
 Db 612 yAsnAlaValAspValArgSerLeuMetGlnValGlnGluThrHisAlaSerSerLeuG1 632
 QY 2583 TTAATTCGAGGATATGGGTTTCTGGAGTTTCGAATTTCTTCATCATGACCGCGATGC 2642
 Db 632 nThrAspArgGlyLeuTrpIleAspGlyIleGlyAenPhePheHisValSerAlaSerG1 652
 QY 2643 TTTAGGTCCAGGATATCGTATATATAGTGGGGTTATCTCTAGGACCAACTCTACTT 2702
 Db 652 uAspAenIleArgTyArgHisAenSerGlyTyTyValLeuSerValAenAenGluI1 672
 QY 2703 TGGATCATCGATGTTT---GGTCTAGCATTTACGAAGTATTTGGTAGATCTAAAGATTA 2759
 Db 672 eThrProLysHisTyThrSerMetAlaPheSerGlnLeuPheSerArgAspLysAspTy 692

QY 2760 TGTAAGTGTGCTTCCCAATCATGCTTGCATAGGATCCGTTTATCTATCTTACCCAAACA 2819
 Db 692 rAlaValSerAenGluTyArg-----MetTyLeu----- 703
 QY 2820 AGCTTTATGTGATCTCTATTTGTC-----GGAGATCGCTTTATCCG 2861
 Db 704 -----GlySerTyLeuTyArgGlnTyThrThrSerLeuGlyAenIlePheArgTy 720
 QY 2862 TGCTAGC----- 2868
 Db 720 rAlaSerArgAenProAenValAenValGlyIleLeuSerArgArgPheLeuGlnAenPr 740
 QY 2869 -----TACGGGTTTGGGAATCAGCATATGAAACCTC 2900
 Db 740 oLeuMetIlePheHisPheLeuCyAlaTyGlyHisAlaThrAenAspMetLysThrAs 760
 QY 2901 ATATCATTTTGCAGAGGAGGATGTTGCTGGGATATAATACTGCTGCTGGCTGGAGAT 2960
 Db 760 pTyAlaAenPheProMetValLysAenSerTrpArgAenAenCysTrpAlaIleGluCy 780
 QY 2961 TGGAGCGGATTTACCGATTTGATTTACTCCATTAAGCTCTATTGGAATGAGTTGCTGC 3020
 Db 780 sGlyGlySerMetProLeuLeuValPheGluAenGlyArgLeuPheGlnGlyAlaIlePr 800
 QY 3021 TTTTCTGCAAGCTGAGTTTCTTATGCTCCGATCATGAATCTTTTACAGAGGAAGCGCATCA 3080
 Db 800 oPheMetLysLeuGlnLeuValTyAlaTyArgGlnGlyAspPheLysGluThrThrAlaAs 820
 QY 3081 AGCTCGGCGATTTCAAGAGCGGACATCTCTCTAAATCTATCAGTTCTCTGTTGGAGTGA 3140
 Db 820 pGlyArgArgPheSerAenGlySerLeuThrSerIleSerValProLeuGlyIleArgPh 840
 QY 3141 TGATCGATGTTCTAGTACACATCTCTAATAAATAGCTTTATGGCGCTTATATCTGTGA 3200
 Db 840 eGluLysLeuAlaLeuSerGlnAspValLeuTyArgPheSerPheSerTyIleProAs 860
 QY 3201 TGCTTATCCACCATCTCTGCTAGTACAGACAAGCTCTCTATCCATCAAGAGACATGGAC 3260
 Db 860 pIlePheArgLysAspProSerCysGluAlaAlaLeuValIleSerGlyAspSerTrpLe 880
 QY 3261 AACAGATGCTTTTCAATTTAGCAAGCATGGAGTTGGTTAGGAGTCTATGTATGCTTC 3320
 Db 880 uValProAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyRH1 900
 QY 3321 TCTAACAGTATATAGATATATGCGCCATCGAAGATATAGTATCGAGATCTTCTCG 3380
 Db 900 sPheAenAspTyThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaAr 920
 QY 3381 AGGCTATGTTTGGTGCAGGAAGTAGAGTCCGGTTC 3417
 Db 920 gAenTyAenIleAenCysGlySerLysPheArgPhe 932
 RESULT 38
 US-09-198-452A-15
 ; Sequence 15 Application US/09198452A
 ; Patent No. 655294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 15
 ; LENGTH: 922
 ; TYPE: PR1
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: 1...922

! OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-15

Alignment Scores:

Pred. No.: 3,03e-80 Length: 922
Score: 964.00 Matches: 304
Percent Similarity: 41.4% Conservative: 149
Best Local Similarity: 27.8% Mismatches: 387
Query Match: 12.2% Indels: 253
DB: 2 Gaps: 35

US-10-701-844-1 (1-4435) x US-09-198-452A-15 (1-922)

```

QY 375 TTGTTGTGATGCAAAAGCTCTTTCATAAGTCTTCTTCTTCAAT-----GAT 419
DB 5 LeuCysGly-----PheProLeuValPheSerPheThrLeuLeuSerValPheAsp 21
QY 420 TCTAGCTTATCTTCTGCTCTTTAAATGGGGGGGATATGCAGCAGAAATCATGTTCC 479
DB 22 ThrSerLeu-SerAlaThrThrIleSer-----LeuThrPr 33
QY 480 TCAAGGAATTTACGATGGGAGACGTTAACTGTATCATTTCCCTATACTGTATTATAGGAGA 539
DB 33 oGluAspSerPheHisGlyAspSerGlnAsnAlaGluArgSerTyrAsnVal----- 50
QY 540 TCCGAGTGGGACTACTGTTTTTTCGAGGAGAGTTAACTAAATAATCTTGACAATTC 599
DB 51 -GlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSerAsnValAspAsnSe 70
QY 600 TATTGACGCTTTCCTTTAAAGTTGTTTGGGAACCTATTAGGAGTTTCTGTTTATAGG 659
DB 70 r-----AlaLeuAsnLysAlaCysPhe***ValThrSerGlySerValThrPheAlaGI 88
QY 660 GAGAGGACACTGTTGACTTTCGAGAACATACGAGCTTCTCA-----AATGGGGCAGC 713
DB 88 YasnHisHisGly***TyrPheAsnAsnIleSerSerGlyThrThrLysGluGlyAlaVa 108
QY 714 TCTAAGTAATAGCGTGTGATGAGTGTCTTACTATGAGGTTTAAAGAAATATCTTT 773
DB 108 lLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPheSerThrLeuSerPh 128
QY 774 TTCCAATTGCAATTCATTACTTGGCTGCTGCTGCTCAACGACTAATAGGTTAGCA 833
DB 128 e-----AsnGI 130
QY 834 GACTCCGACGACACATCTACACCGCTAATGGTACTATTATTCTTAAACAGACTCTTT 893
DB 130 nSerPro-----GlyAspIleLysGluGlnGlyCysLeuTyrSerLysAsnAlaLeuMe 148
QY 894 GTTACTCAATAATAGAAATCTCTATTCTATAGTAATTTAGTCTCTGGAGATGGGGAGC 953
DB 148 tLeuLeuAsnAsnTyrValValArgPheGluGlnAsnGlnSerLysThrLysGlyGlyAl 168
QY 954 TATAGATGCTAGAGCTTAACGGTTCAAGGAATTAGCAAGCTTTGTGCTTCCAGAGAAA 1013
DB 168 aIleSerGlyAlaAsnValThrIleValGlyAsnTyrAspSerValSerPheTyrGlnAs 188
QY 1014 TACTGCTCAAGCTGATGGGGAGCTGTCAAGTAGCTGCTCAACAGTTCTCTGCTATGCTAA 1073
DB 188 nAlaAlaThrPheGlyGly-----AlaIleHisE 198
QY 1074 CGAGGCTCTATTGCTTTTGTAGCGAATGTTCCAGAGGTAAGAGGGGAGGATGCTGTC 1133
DB 198 rSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArg----- 212
QY 1134 TGTTGAGGATGGGACGAGGAGTGTCTATCTACTTCTCAACAGAGATCCAGTAGTAG 1193
DB 212 ----- 212
QY 1194 TTTTTCAGAAATACTGCGGTAGAGTTTGTATGGGAACGTAGCCCGAGTAGGAGGAGAT 1253
DB 213 -PheAlaGlnAsnThrAlaLys-----AsnGlySer-----GlyGlyAlaLe 226

```

```

QY 1254 TTACTCTACGGGAACGTTGCTTCTCTGAATAATAGGAAACCTTGTTCTCTCAACAATGT 1313
DB 226 uTyrSerAspGlyAspIleAspIleAspGlnAsnAlaTyrValLeuPheArgGluAsnGI 246
QY 1314 TGCTTCTCTGTTTACATTTGCTGTAAAGCAACCAAGAGTGCAGGCTTCTTAATACGAG 1373
DB 246 uAlaLeuThrThrAlaIle----- 252
QY 1374 TAATAAATTTACGAGATGGAGAGCTATCTTCTGTGAAGAATGTTGCGCAAGCAGGATCAA 1433
DB 253 -----GlyLysGlyGlyAlaValCysCys-----LeuPr 262
QY 1434 TAACTCTGGATCA-----GTTCTCTTTCATGGAGGAGGAGTAGTTT 1475
DB 262 oThrSerGlySerSerThrProValProIleValThrPheSerAspAsnLysGlnLeuVa 282
QY 1476 CTTTAGTAGCAATGATGCTGCGGAAAGGGGAGCTATTATTCGCAAAAAGCTCTCGGT 1535
DB 282 lPheGluArgAsnHisSerIleMetGlyGlyAlaIleTyrAlaArgLysLeuSerIl 302
QY 1536 TGCTAACTGTGGCCTGTACAAATTTTAAAGAAATATCGCT-----AATGA 1580
DB 302 eSerSerGlyGlyProThrLeuPheIleAsnAsnIleSerTyrAlaAsnSerGlnAsnLe 322
QY 1581 TGCTGGAGCGGATTTATTAGGAGAACTCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGA 1640
DB 322 uGlyGlyAlaIleAlaIleAspThrGlyGlyIleSerLeuSerAlaGlyLysGlyTh 342
QY 1641 TATTATTTCGATCGGAATCTTAAAGAAACAGCAAGAGAAATGCTGCGGATGTTAAATGG 1700
DB 342 rIleThrPheGlnGlyAsn-----ArgThrSerLeuProPheLeuAsnGI 357
QY 1701 CGTAACTGTGTCTCACAAAGCCATTTTCATGCGGATCGGAGGAAATAACGACATTAAG 1760
DB 357 YIleHisLeuLeuGlnAsnAla-----LysPheLeuLysLeuGI 370
QY 1761 AGCTAAAGCAGGCGATCGATTTCTTTTAATGATCCCATCGAGATGGCAACGGAATAA 1820
DB 370 nAlaArgAsnGlyTyrSerIleGlyPheTyrAspProIle-----Th 384
QY 1821 CCAGCCAGCGCAGTCTTCCAAACTTCTAAATAATTAAAC-----GATGGTGAAGGATA 1871
DB 384 rSerGluAlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLysGluTy 404
QY 1872 CACAGGGGATATGTTTTGCT-----AATGGAAG 1901
DB 404 rThrGlyThrIleLeuPheSerGlyGluLysSerLeuAlaAsnAspProArgAspPheLy 424
QY 1902 CAGTACTTTGTACCAAAATGTTACGATAGACGAGGAGGATTTGTTCTGTTGAAAAGGC 1961
DB 424 sSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyrLeuValIleLysGluGlyAl 444
QY 1962 AAAATTATCAGTGAATCTCTAAGTCACAGAGTGGAGT---CTGTATATGGAAGCTGG 2018
DB 444 aGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeuAspLeuGI 464
QY 2019 GAGTACATGGGATTTTGTAACTCCACAAACCAACAGCCTCTCTCGCGCTTAATCAGTT 2078
DB 464 yThr-----LysLeuIleAlaSerLysGluAs 473
QY 2079 GATCAGCTTTCCAAATCTGCATTTGCTCTTCTTCTTTGTTAGCAACAATGCGATTAC 2138
DB 473 pIleAlaIleThrGlyLeuAlaIleAspIleAspSerLeu----- 486
QY 2139 GAATCTCTTACCATCTCCAGCGCAAGATTTCTCATCTCGAGTCAATTGGTAGCACAA 2198
DB 487 -----SerSerSerThrAlaAlaValIleLysAlaAsnTh 499
QY 2199 TGCTGGT-----TCTGTTACAAATTAGTGGCGCTATCTTT----- 2232
DB 499 rAlaAsnLysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAl 519
QY 2233 -TTTGAGGATTTGGATGATACAGCTTATGATAGGTATGATTGGCTAGGTTCTTAATCAAAA 2291

```



```

QY 2797 TCCGTTTATCTATCCCAACAGCTTTATGTGGA-----TCCTATTGTTCCGAGAT 2850
DB 637 SerLeuArgLeuGlnHisAspAlaSerLeuTyrSerValValSerLeuLeuGlyGlu 656
QY 2851 GCGTTTATCCGT----- 2862
DB 657 GlyGlyLeuArgGluLeuLeuProTyrValSerLeuThrLeuProCysSerPheTyr 676
QY 2863 -----GCTAGTACGGTTTGGGAATCAGCATATGAAACCTCATAT----- 2904
DB 677 GlyGlnLeuSerTyrGlyHisThrAspHisArgMetLeuThrGluSerLeuProPro 696
QY 2905 -----ACATTTGCAGAGAGAGCGATGTTCTGTCGGAATAAATCTGTCGTCGAGAG 2958
DB 697 ProProThrLeuSerThrAsp---HisThrSerTrpGlyTyrValTrpAlaGlyGlu 715
QY 2959 ATTGAGCGGGATTACCGATTGTAATCTCCATCTTAAGCTCTATTGTAATGAGTTGCGT 3018
DB 716 LeuGlyThrArgValAlaValGluAlaThrSerGlyArgGlyPhePheGlnGluTyrThr 735
QY 3019 CTTTTCGTCAGAGCTGAGTTTCTTATGCGCATCATGAAATCTTTTACAGAGAGCGCAT 3078
DB 736 ProPheValLysValGlnAlaValTyrAlaArgGlnAspSerPheValGluLeuGlyAla 755
QY 3079 CAAGCTCGGCATTCAGAGCGGACATCTCCTAAATCTATCATGTTCTGTCGAGTGAAG 3138
DB 756 IleSerArgAspPheSerAspSerHisLeuTyrAsnLeuAlaIleProLeuGlyIleLys 775
QY 3139 TTTGATCGATGTTCTAGTACACATCTTAATAATATAGCTTTTATGCGGCTTATCTGT 3198
DB 776 LeuGluLysArgPheAla-----GluGlnTyrTyrHisValAlaMetTyrSerPro 793
QY 3199 GATGCTTATCGACCATCTCTGTACTGAGACACCGCTCTATCCATCAAGAGCATGG 3258
DB 794 AspValCysArgSerAsnProLysCysThrThrThrLeuLeuSerAsnGlnGlySerTrp 813
QY 3259 ACAACAGATGCTTTTCATTTAGCAGACATGAGGTTGCTGTTAGAGGATCTGTATGCT 3318
DB 814 LysThrLysGlySerAsnLeuAlaArgGlnAlaGlyIleValGlnAlaSerGlyPheArg 833
QY 3319 TCTCTAACAAATATAGAGTATATGCGCATGGAAGATATGAGTATGAGATGCTTCT 3378
DB 834 SerLeuGlyAlaAlaAlaGluLeuPheGlyAsnPheGlyPheGluTrpArgGlySerSer 853
QY 3379 CGAGGCTATGTTGATGTCAGGAGTAGAGTCCGGTTC 3417
DB 854 ArgSerTyrAsnValAspAlaGlySerLysLeuLysPhe 866

RESULT 40
US-09-198-452A-466
; Sequence 466, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; thereof and uses thereof, in particular for the diagnosis, prevention
; and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 466
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-466

Alignment Scores:
Pred. No.: 1,44e-67 Length: 1132
Score: 828.00 Matches: 291
Percent Similarity: 37.1% Conservative: 117
Best Local Similarity: 26.5% Mismatches: 327

```

```

Query Match: 10.5% Indels: 365
DB: 2 Gaps: 35
US-10-701-844-1 (1-4435) x US-09-198-452A-466 (1-1132)

QY 493 GATGGGAGAGCTTAACCTGATCATTTCCCTATATCTGTTATAGGAGATCCGAGTGGACT 552
DB 36 AsnGlySerSerGlySerAlaAlaPheThrAlaLysGluThrSerAspAlaSerGlyThr 55
QY 553 ACTGTTTTTCTGCAGGAGAGTTACATTAATAAATCTTGACAAATCTTGACAAATCTTATTCAGCTTGG 612
DB 56 ThrTyrThrLeuThrSerAspValSerIleThrAsnVal---SerAlaIleThrProAla 74
QY 613 CCTTTAAGTTGTTTGGGAATCTATTAGGAGCTTTTACTGTTTATAGGAGAGACACTCG 672
DB 75 AspLysSerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSer 94
QY 673 TTGACTTTTCGAGAACATACCGACTTCTCAAAATGCGGAGCTCTTAAGTAAATAGCGCTGT 732
DB 95 LeuValLeuGlnThrIleAlaLeuThrHisAspGlyAlaAlaIleAsnAsnThrAsnThr 114
QY 733 GATGACTGTTTACTATTGAGGGTTTAAAGAAATATCTTTTCCAAATTCGAATTCATTA 792
DB 115 -----AlaLeuSerPheSerGlyPheSerSerLeu 124
QY 793 CTT----- 795
DB 125 LeuIleAspSerAlaProAlaThrGlyThrSerGlyGlyLysGlyAlaIleCysValThr 144
QY 795 ----- 795
DB 145 AsnThrGluGlyThrAlaThrPheThrAspAsnAlaSerValThrLeuGlnLysAsn 164
QY 795 ----- 795
DB 165 ThrSerGluLysAspGlyAlaAlaValSerAlaTyrSerIleAspLeuAlaLysThrThr 184
QY 796 -----GCCGTACTGCTGCT 810
DB 185 ThrAlaAlaLeuLeuAspGlnAsnThrSerThrLysAsnGlyGlyAlaLeuCysSerThr 204
QY 811 GCAACGACT-----AATAGGCTAGCCAGACT---CCGACGACCAACATCTACA 855
DB 205 AlaAsnThrThrValGlnGlyAsnSerGlyThrValThrPheSerSerAsnThrAlaThr 224
QY 856 CGCTCTAAGCTGACTATTATCT----- 879
DB 225 AspLysGlyGlyIleTyrSerLysGluLysAspSerThrLeuAspAlaAsnThrGly 244
QY 879 ----- 879
DB 245 ValValThrPheLysSerAsnThrAlaLysThrGlyGlyAlaTrpSerSerAspAsn 264
QY 879 ----- 879
DB 265 LeuAlaLeuThrGlyAsnThrGlnValLeuPheGlnGluAsnLysThrThrGlySerAla 284
QY 879 ----- 879
DB 285 AlaGlnAlaAsnAsnProGluGlyCysGlyAlaIleCysCysTyrLeuAlaThrAla 304
QY 880 -----AAACACAGATCTTTGTTTACTCAATATGAGAAGTTCTCATTTCTATAGTAATTTA 933
DB 305 ThrAspLysThrGlyLeuAlaIleSerGlnAsnGlnGluMetSerPheThrSerAsnThr 324
QY 934 GTCTCTGAGATGGGGAGCTATAGATGCTTAAGAGCTTAACGGTTCAAGAAATTAGCAAG 993
DB 325 ThrThrAlaAsnGlyGlyAlaIleTyrAlaThrLysCysThrLeuAspGlyAsnThrThr 344
QY 994 CTTTGTGCTTCCAGAAATATCTGCTCAAGCTGAT---CGGGAGCTTCTCAAGTAGTC 1050
DB 345 Leu---ThrPheAspGlnAsnThrAlaThrAlaGlyCysGlyGlyAlaIleTyrThrGlu 363

```


Db 1041 LysAspTyrValValSerAspIleLysSerGlnValTyrAlaGlySer-----LeuCys 1058
Qy 2812 ACCCAACAGCTTTATGTGGATCCTATTTGTTCGGAGATGGCTTTATCCGTGCTAGCT 2869
Db 1059 AlaGlnSerSer-TyrValIleProLeuHisSerSerLeuArgArgHisValLeuSer 1077

Search completed: May 13, 2006, 10:40:44
Job time : 424 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:20:04 ; Search time 48 seconds
(without alignments)
1743.079 Million cell updates/sec

Title: US-10-701-844-2

Perfect score: 5267

Sequence: 1 MQTSFKFPLSMILAYSCCS.....YEYRDASRGVGLSAGSRVRF 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq.length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/PCRTUS COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/FTUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5267	100.0	1012	2	US-09-612-402B-2
2	5267	100.0	1012	2	US-09-542-520-2
3	5116	97.1	984	2	US-09-612-402B-43
4	5106.5	97.0	1013	2	US-09-612-402B-15
5	5106.5	97.0	1013	2	US-09-612-402B-16
6	5106.5	97.0	1013	2	US-09-542-520-15
7	5106.5	97.0	1013	2	US-09-542-520-16
8	5090	96.6	1006	2	US-09-556-877-190
9	5090	96.6	1006	2	US-09-620-412C-190
10	5090	96.6	1006	2	US-09-598-419-190
11	5084	96.5	982	2	US-09-556-877-176
12	5084	96.5	982	2	US-09-620-412C-176
13	5084	96.5	982	2	US-09-598-419-176
14	3336.5	63.3	670	2	US-10-197-220-169
15	2552	48.5	505	2	US-09-612-402B-17
16	2547	48.4	505	2	US-09-542-520-17
17	2350	44.6	458	2	US-09-612-402B-36
18	2350	44.6	458	2	US-09-542-520-36
19	1735	32.9	325	2	US-09-612-402B-37
20	1735	32.9	325	2	US-09-542-520-37
21	1365.5	25.9	999	2	US-09-438-185A-455
22	1359.5	25.8	973	2	US-09-430-723-2
23	1128	21.4	928	2	US-09-428-122-2
24	1090	20.7	949	2	US-09-198-452A-478
25	1090	20.7	967	2	US-09-438-185A-453
26	1083	20.6	930	2	US-09-198-452A-470
27	1079	20.5	938	2	US-09-438-185A-448

28	1041	19.8	947	2	US-09-438-185A-447	Sequence 447, App
29	1021	19.4	937	2	US-09-438-185A-449	Sequence 449, App
30	994.5	18.9	927	2	US-09-198-452A-472	Sequence 472, App
31	985.5	18.7	1414	2	US-09-438-185A-446	Sequence 446, App
32	980.5	18.6	780	2	US-09-438-185A-17	Sequence 17, Appl
33	967	18.4	932	2	US-09-438-185A-6	Sequence 6, Appl
34	958	18.2	922	2	US-09-198-452A-15	Sequence 15, Appl
35	939.5	17.8	866	2	US-09-438-185A-15	Sequence 15, Appl
36	821.5	15.6	1132	2	US-09-198-452A-466	Sequence 466, App
37	808	15.3	634	2	US-09-438-185A-451	Sequence 451, App
38	804	15.3	643	2	US-09-198-452A-474	Sequence 474, App
39	769.5	14.6	880	2	US-09-556-877-175	Sequence 175, App
40	769.5	14.6	880	2	US-09-620-412C-175	Sequence 175, App
41	769.5	14.6	880	2	US-09-598-419-175	Sequence 175, App
42	762.5	14.5	866	2	US-09-556-877-189	Sequence 189, App
43	762.5	14.5	866	2	US-09-620-412C-189	Sequence 189, App
44	762.5	14.5	866	2	US-09-598-419-189	Sequence 189, App
45	725	13.8	597	2	US-09-198-452A-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1

US-09-612-402B-2

; Sequence 2, Application US/09612402B

; Patent No. 6642023

; GENERAL INFORMATION:

; APPLICANT: Jackson, W. James

; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

; Patent No. 6642023

; FILE REFERENCE: 7969-086-999

; CURRENT APPLICATION NUMBER: US/09/612,402B

; PRIOR FILING DATE: 2000-07-06

; PRIOR APPLICATION NUMBER: 08/942,596

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1012

; TYPE: PRT

; ORGANISM: Chlamydia sp.

US-09-612-402B-2

Query Match	100.0%;	Score 5267;	DB 2;	Length 1012;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1012;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MQTSFKFPLSMILAYSCCSLNGGGYAAETWVPGIYDGETLTVSPYTVIGDPSTTVF	60	
Db	1	MQTSFKFPLSMILAYSCCSLNGGGYAAETWVPGIYDGETLTVSPYTVIGDPSTTVF	60	
Qy	61	SAGELTKNLDNSIAALPLSCFGLNLSFTVLGRGHSLTPEINTRTNGAALSNSAADGL	120	
Db	61	SAGELTKNLDNSIAALPLSCFGLNLSFTVLGRGHSLTPEINTRTNGAALSNSAADGL	120	
Qy	121	FTTEGPKELFSNCNSLLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF	180	
Db	121	FTTEGPKELFSNCNSLLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKFSF	180	
Qy	181	YSNLVSGDGAIDAKSITVQGISKLVQFQENTQAADGACQVVTFSFMANEAPIAFVAN	240	
Db	181	YSNLVSGDGAIDAKSITVQGISKLVQFQENTQAADGACQVVTFSFMANEAPIAFVAN	240	
Qy	241	VAGVRGGGIAAVDGGQGVSSSTSTEDPVVSFRNTAVEFDGNVARYGGGIYSYGNVAF	300	
Db	241	VAGVRGGGIAAVDGGQGVSSSTSTEDPVVSFRNTAVEFDGNVARYGGGIYSYGNVAF	300	
Qy	301	NGKTLPLNNVAPVYIAAKQPTSGQASNTSNNGYDGGAIIFCKNGAAGSNSSGVSPDG	360	
Db	301	NGKTLPLNNVAPVYIAAKQPTSGQASNTSNNGYDGGAIIFCKNGAAGSNSSGVSPDG	360	

QY 361 EGVVFPSSNVAAGKGAIIYAKKLSVANGCPVQFLRNANDGGAIYLGSGELSLADYGD 420
DB |||||
QY 361 EGVVFPSSNVAAGKGAIIYAKKLSVANGCPVQFLRNANDGGAIYLGSGELSLADYGD 420
DB |||||
QY 421 IIFDGNLKRKTAKENAADVNGVTVSSQAISMGGKITTILRAKAGHQIILFNDPIEMANGNN 480
DB |||||
QY 421 IIFDGNLKRKTAKENAADVNGVTVSSQAISMGGKITTILRAKAGHQIILFNDPIEMANGNN 480
DB |||||
QY 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQRIIVLRKAKLSVNSLSQTGGS 540
DB |||||
QY 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQRIIVLRKAKLSVNSLSQTGGS 540
DB |||||
QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSISLANNVNTNPPPAQDSHPA 600
DB |||||
QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSISLANNVNTNPPPAQDSHPA 600
DB |||||
QY 601 VIGSTTAGSVTISGPIFPEDLDDTDAYDRYDMLGSKINVLKQLGTQKPPANAPSDLTIG 660
DB |||||
QY 601 VIGSTTAGSVTISGPIFPEDLDDTDAYDRYDMLGSKINVLKQLGTQKPPANAPSDLTIG 660
DB |||||
QY 661 NEMPKYGVQGSWKLAWDPNTANGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDIR 720
DB |||||
QY 661 NEMPKYGVQGSWKLAWDPNTANGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDIR 720
DB |||||
QY 721 SAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDALGQGYRISGYSLGANSYFGSSMFG 780
DB |||||
QY 721 SAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDALGQGYRISGYSLGANSYFGSSMFG 780
DB |||||
QY 781 LAFTVEFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKTS 840
DB |||||
QY 781 LAFTVEFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKTS 840
DB |||||
QY 841 YTFABESDVRWNNCLAGEIGAGLPVITPSPKLYLNELRPFVQAFSFIADHESFTEEGDQ 900
DB |||||
QY 841 YTFABESDVRWNNCLAGEIGAGLPVITPSPKLYLNELRPFVQAFSFIADHESFTEEGDQ 900
DB |||||
QY 901 ARAFSGHLLNLSPVGVKFDRCSTHFNKYSFMAAYICDAVRTISGTTLLSHQETWT 960
DB |||||
QY 901 ARAFSGHLLNLSPVGVKFDRCSTHFNKYSFMAAYICDAVRTISGTTLLSHQETWT 960
DB |||||
QY 961 TDAFHARHGTVVVRGSMYASLTNSIEVYGHGREYRDASRGYGLSAGSRVRF 1012
DB |||||
QY 961 TDAFHARHGTVVVRGSMYASLTNSIEVYGHGREYRDASRGYGLSAGSRVRF 1012
DB |||||

RESULT 2

US-09-542-520-2
; Sequence 2, Application US/09542520
; Patent No. 6867843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-542-520-2

Query Match 100.0%; Score 5267; DB 2; Length 1012;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTSFHKFFLSMILAYSCSLNGGGYAAEIMVPPQGIYDGETLTVSFPYTVIGDPSGTTVVF 60
DB 1 MQTSFHKFFLSMILAYSCSLNGGGYAAEIMVPPQGIYDGETLTVSFPYTVIGDPSGTTVVF 60

QY 61 SAGELTLKXLDNSIAALPLSCFQNLGSLFTVLGRGHSITPENIRITSTNGAALSNSAADGL 120
DB |||||
QY 61 SAGELTLKXLDNSIAALPLSCFQNLGSLFTVLGRGHSITPENIRITSTNGAALSNSAADGL 120
DB |||||
QY 121 FTIEGFKELSPSNCNLSLAVLPAATTNKGSTPTTTSTPSNGTTSYKTDLLLLNNEKFSF 180
DB |||||
QY 121 FTIEGFKELSPSNCNLSLAVLPAATTNKGSTPTTTSTPSNGTTSYKTDLLLLNNEKFSF 180
DB |||||
QY 181 YSNLVSGDGGDAIDAKSLTVQGISKLCPVQENTAQADGGACQVTSFSSAMANEAPLAFVAN 240
DB |||||
QY 181 YSNLVSGDGGDAIDAKSLTVQGISKLCPVQENTAQADGGACQVTSFSSAMANEAPLAFVAN 240
DB |||||
QY 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVEFDGNVARVCGGIIYSYGNVAPL 300
DB |||||
QY 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVEFDGNVARVCGGIIYSYGNVAPL 300
DB |||||
QY 301 NNGKTLPLNNVASPVYIAAKOPTSQASNTSNNTYDGGAIIFCKNGAQAGSNNSGSVSFDG 360
DB |||||
QY 301 NNGKTLPLNNVASPVYIAAKOPTSQASNTSNNTYDGGAIIFCKNGAQAGSNNSGSVSFDG 360
DB |||||
QY 361 EGVVFPSSNVAAGKGAIIYAKKLSVANGCPVQFLRNANDGGAIYLGSGELSLADYGD 420
DB |||||
QY 421 IIFDGNLKRKTAKENAADVNGVTVSSQAISMGGKITTILRAKAGHQIILFNDPIEMANGNN 480
DB |||||
QY 421 IIFDGNLKRKTAKENAADVNGVTVSSQAISMGGKITTILRAKAGHQIILFNDPIEMANGNN 480
DB |||||
QY 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQRIIVLRKAKLSVNSLSQTGGS 540
DB |||||
QY 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQRIIVLRKAKLSVNSLSQTGGS 540
DB |||||
QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSISLANNVNTNPPPAQDSHPA 600
DB |||||
QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSISLANNVNTNPPPAQDSHPA 600
DB |||||
QY 601 VIGSTTAGSVTISGPIFPEDLDDTDAYDRYDMLGSKINVLKQLGTQKPPANAPSDLTIG 660
DB |||||
QY 601 VIGSTTAGSVTISGPIFPEDLDDTDAYDRYDMLGSKINVLKQLGTQKPPANAPSDLTIG 660
DB |||||
QY 661 NEMPKYGVQGSWKLAWDPNTANGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDIR 720
DB |||||
QY 661 NEMPKYGVQGSWKLAWDPNTANGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDIR 720
DB |||||
QY 721 SAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDALGQGYRISGYSLGANSYFGSSMFG 780
DB |||||
QY 721 SAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDALGQGYRISGYSLGANSYFGSSMFG 780
DB |||||
QY 781 LAFTVEFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKTS 840
DB |||||
QY 781 LAFTVEFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKTS 840
DB |||||
QY 841 YTFABESDVRWNNCLAGEIGAGLPVITPSPKLYLNELRPFVQAFSFIADHESFTEEGDQ 900
DB |||||
QY 841 YTFABESDVRWNNCLAGEIGAGLPVITPSPKLYLNELRPFVQAFSFIADHESFTEEGDQ 900
DB |||||
QY 901 ARAFSGHLLNLSPVGVKFDRCSTHFNKYSFMAAYICDAVRTISGTTLLSHQETWT 960
DB |||||
QY 901 ARAFSGHLLNLSPVGVKFDRCSTHFNKYSFMAAYICDAVRTISGTTLLSHQETWT 960
DB |||||
QY 961 TDAFHARHGTVVVRGSMYASLTNSIEVYGHGREYRDASRGYGLSAGSRVRF 1012
DB |||||
QY 961 TDAFHARHGTVVVRGSMYASLTNSIEVYGHGREYRDASRGYGLSAGSRVRF 1012
DB |||||

RESULT 3

US-09-612-402B-43
; Sequence 43, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John

; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PR1
; ORGANISM: Chlamydia sp.
US-09-612-402B-43

Query Match 97.1%; Score 5116; DB 2; Length 984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 EIMVPGIYDGETLTIVSPYTVIGDPSGTTVFSAGELTLKNDNSIAALPLSCFGLNLS 88
DB 1 EIMVPGIYDGETLTIVSPYTVIGDPSGTTVFSAGELTLKNDNSIAALPLSCFGLNLS 60
QY 89 FTVLGRGHSLTPEINRTSTNGAALSNSAADGLFTIEGFKELSPNSCNLSIAVLPAATTNK 148
DB 61 FTVLGRGHSLTPEINRTSTNGAALSNSAADGLFTIEGFKELSPNSCNLSIAVLPAATTNK 120
QY 149 GSOTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLNLSVSGDGAIDAKSLTVQGISKLCVF 208
DB 121 GSOTPTTSTPSNGTIYSKTDLLLNNEKFSFYSLNLSVSGDGAIDAKSLTVQGISKLCVF 180
QY 209 QENTAQDGCACQVTSFSAANEAPAFVANVAGVGGGIAAVQDQGGVSSSTEDP 268
DB 181 QENTAQDGCACQVTSFSAANEAPAFVANVAGVGGGIAAVQDQGGVSSSTEDP 240
QY 269 VVSFSPRTAVEFGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQAS 328
DB 241 VVSFSPRTAVEFGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQAS 300
QY 329 NTSNNYGDGGAIFCKNGAAGSNNSSVSFDGSGVFPFSSNVAAGKGAIYAKKLSVANC 388
DB 301 NTSNNYGDGGAIFCKNGAAGSNNSSVSFDGSGVFPFSSNVAAGKGAIYAKKLSVANC 360
QY 389 GPVQFLRNANDGGAIVLGSSELSSADYGDIIFDGNLKRKTAKENADVNGTVSSQAI 448
DB 361 GPVQFLRNANDGGAIVLGSSELSSADYGDIIFDGNLKRKTAKENADVNGTVSSQAI 420
QY 449 SMGSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTDIVFANGSS 508
DB 421 SMGSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTDIVFANGSS 480
QY 509 TLQNVTVIEQRIVLRKAKLSVNSLSQTGSLYMEAGSTWDFVTPQPPQPPAANQLIT 568
DB 481 TLQNVTVIEQRIVLRKAKLSVNSLSQTGSLYMEAGSTWDFVTPQPPQPPAANQLIT 540
QY 569 LSNLHLSLSLLANNAVTPNTPPAQDSHPAVTGSTAGSVTISGPIFFEDLDDTAYDR 628
DB 541 LSNLHLSLSLLANNAVTPNTPPAQDSHPAVTGSTAGSVTISGPIFFEDLDDTAYDR 600
QY 629 YDWLGSNQKINVLKQLGTGPPANAPSDLTILGNEMPKYVQGSWKLAWDPNTANNPYTL 688
DB 601 YDWLGSNQKINVLKQLGTGPPANAPSDLTILGNEMPKYVQGSWKLAWDPNTANNPYTL 660
QY 689 KATWTKYGNPGERVASLPNSLWGSTLDIRSAHSAIQASVDRGSCYRGLWVSGVSNFF 748
DB 661 KATWTKYGNPGERVASLPNSLWGSTLDIRSAHSAIQASVDRGSCYRGLWVSGVSNFF 720
QY 749 YHREDALGQGYRISGGYSLGANSYFGSSMFGLAFTVEFGRSKDYVVCRRSNHACIGSVY 808
DB 721 YHREDALGQGYRISGGYSLGANSYFGSSMFGLAFTVEFGRSKDYVVCRRSNHACIGSVY 780
QY 809 LSTQOALCGSYLFGDAFIRASYGFGNOHMKTSYTFABESDVRWNNCLAGIGAGLPVI 869

DB 781 LSTQOALCGSYLFGDAFIRASYGFGNOHMKTSYTFABESDVRWNNCLAGIGAGLPVI 840
QY 869 TPSKLYLNLRLRPVQAEFSYADHESFTTEGDOARAFKSHLNLNLSVPVGVKFDRCSSTHP 928
DB 841 TPSKLYLNLRLRPVQAEFSYADHESFTTEGDOARAFKSHLNLNLSVPVGVKFDRCSSTHP 900
QY 929 NKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGCVVVRGSMYASLTSNIEVY 988
DB 901 NKYSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGCVVVRGSMYASLTSNIEVY 960
QY 989 GHGREYRDASRGYGLSAGSRVF 1012
DB 961 GHGREYRDASRGYGLSAGSRVF 984

RESULT 4

US-09-612-402B-15
; Sequence 15, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PR1
; ORGANISM: Chlamydia sp.
US-09-612-402B-15

Query Match 97.0%; Score 5106.5; DB 2; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;
QY 1 MOTSFHKFFLSMLIAYSCCLNGGGYAAEIMVPGIYDGETLTIVSPYTVIGDPSGTTVF 60
DB 1 MOTSFHKFFLSMLIAYSCCLNGGGYAAEIMVPGIYDGETLTIVSPYTVIGDPSGTTVF 60
QY 61 SAGELTLKNDNSIAALPLSCFGLNLSFTVLGRGHSLTPEINRTSTNGAALSNSAADGL 120
DB 61 SAGELTLKNDNSIAALPLSCFGLNLSFTVLGRGHSLTPEINRTSTNGAALSNSAADGL 120
QY 121 FTIEGFKELSPNSCNLSIAVLPAATTNKSGTSOTPTTSTPSNGTIYSKTDLLLNNEKFSF 180
DB 121 FTIEGFKELSPNSCNLSIAVLPAATTNKSGTSOTPTTSTPSNGTIYSKTDLLLNNEKFSF 180
QY 181 YSNLSVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVTSFSAANEAPAFVAN 240
DB 181 YSNLSVSGDGAIDAKSLTVQGISKLCVFOENTAQAADGGACQVTSFSAANEAPAFVAN 240
QY 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEPDGNVARVGGIYSYGNVAF 300
DB 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEPDGNVARVGGIYSYGNVAF 300
QY 301 NNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGYGGGAIKCKNGAQ-AGSNNSSVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGYGGGAIKCKNGAQ-AGSNNSSVSFD 360
QY 360 GEGVVFSSNVAAGKGAIYAKKLSVANCGPVQFLRNANDGGAIVLGSSELSSADY 419
DB 361 GEGVVFSSNVAAGKGAIYAKKLSVANCGPVQFLRNANDGGAIVLGSSELSSADY 420
QY 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 479
DB 421 DMIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 480

Qy	480	NPQAQSSKLLKINDGEGYTGDIIVFANGSGSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG	533
Db	481	NPQAQSSKLLKINDGEGYTGDIIVFANGSGSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG	540
Qy	540	SLYWEAGSTWDFVTPQPQPPAANOLITLSNHLISLSLLANNAVTPPTNPPAQDSHP	599
Db	541	SLYWEAGSTWDFVTPQPQPPAANQSIITLSNHLISLSLLANNAVTPPTNPPAQDSHP	600
Qy	600	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQKINVLKQLGTKPPANAPSDLT	659
Db	601	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQKIDVLKQLGTPPANAPSDLT	660
Qy	660	GNEMPKYGYQGSWKLANDPNTANNPPTLKATWTKTGYNPGERVASIVPNSLWGSILDI	719
Db	661	GNEMPKYGYQGSWKLANDPNTANNPPTLKATWTKTGYNPGERVASIVPNSLWGSILDI	720
Qy	720	RSAHSAIQASVDGRSYCRGLWVGVSNNFFYHDDRDLGQGYRYISGYSILGANSYFGSSMF	779
Db	721	RSAHSAIQASVDGRSYCRGLWVGVSNNFFYHDDRDLGQGYRYISGYSILGANSYFGSSMF	780
Qy	780	GLATTEVFGRSKDVVCRSNHHACIGSVYLSVTSQALCGSYLFGDAFIRASYGFGNQHKMT	839
Db	781	GLATTEVFGRSKDVVCRSNHHACIGSVYLSVTSQALCGSYVFGDAFIRASYGFGNQHKMT	840
Qy	840	SYTFAESDVRWDDNCLAGEIGAGLPIVITPSKLYLNLRPVQAEFYADHESFTEGD	899
Db	841	SYTFAESDVRWDDNCLAGEIGAGLPIVITPSKLYLNLRPVQAEFYADHESFTEGD	900
Qy	900	QARAFKSGHLLNLSPVGVKFDRCSTHPNKYSFMAAIVICDARTISGTETLLSHQETW	959
Db	901	QARAFKSGHLLNLSPVGVKFDRCSTHPNKYSFMAAIVICDARTISGTETLLSHQETW	960
Qy	960	TTDAFHLARHGCVVRGSGMYASITSNIEVYGHGREYRDASRCYGLSAGSRVRF	1012
Db	961	TTDAFHLARHGCVVRGSGMYASITSNIEVYGHGREYRDTSRGYLSAGSKVRP	1013
RESULT 5			
US-09-612-402B-16			
; Sequence 16, Application US/09612402B			
; Patent No. 6642023			
; GENERAL INFORMATION:			
; APPLICANT: Jackson, W. James			
; APPLICANT: Pace, John			
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof			
; Patent No. 6642023			
; FILE REFERENCE: 7369-086-999			
; CURRENT APPLICATION NUMBER: US/09/612.402B			
; CURRENT FILING DATE: 2000-07-06			
; PRIOR APPLICATION NUMBER: 08/942,596			
; PRIOR FILING DATE: 1997-10-02			
; NUMBER OF SEQ ID NOS: 43			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 16			
; LENGTH: 1013			
; TYPE: PRT			
; ORGANISM: Chlamydia sp.			
US-09-612-402B-16			
Query Match 97.0%; Score 5106.5; DB 2; Length 1013;			
Best Local Similarity 96.7%; Pred. No. 0;			
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1			
Qy	1	MQTSFHKFFLSMILAYSCCSLNGGYYAAEIMVPOQIYDGETLTVSPFTYVIGDPSGTTVF	60
Db	1	MQTSFHKFFLSMILAYSCCSLNGGYYAAEIMVPOQIYDGETLTVSPFTYVIGDPSGTTVF	60
Qy	61	SAGELTTLKLDNSTAALPLSCFGNLIGSFVTLGRGHSITPENIRSTNGAALSNAADGL	120
Db	61	SAGELTTLKLDNSTAALPLSCFGNLIGSFVTLGRGHSITPENIRSTNGAALSNAADGL	120
Qy	121	FTIEGPKELSFSCNCSLILAVLPAATTKNGSGTPTTTSTPSNGTYSKYDILLANNEKFSF	180

[illegible]

RESULT 8

```

US-09-556-877-190
Sequence 190, Application US/09556877
Patent NO. 6432916
GENERAL INFORMATION:
APPLICANT: Probst, Peter
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasar
APPLICANT: Fling, Steve
APPLICANT: Maisonneuve, Jeff
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C5
CURRENT APPLICATION NUMBER: US/09/556,877
CURRENT FILING DATE: 2000-04-19
NUMBER OF SEQ ID NOS: 305
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 190
LENGTH: 1006
TYPE: PRT
ORGANISM: Chlamydia
US-09-556-877-190

```

Query Match	96.6%	Score 5090;	DB 2;	Length 1006;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches	979;	Conservative 2;	Mismatches 1;	Indels 0; Gaps 0
Qy	31	MVPGGIYDGETLTVSPPYTVTVGDPSGTTVP	SAGELTLKXLDNSIAALPLSCF	GNLLGSFT 90
		:	:	:
Db	25	MIPGGIYDGETLTVSPFYTVTVGDPSGTTVP	SAGELTLKXLDNSIAALPLSCF	GNLLGSFT 84
Qy	91	VLGRGHSILTPENIRTSITNGAALSNSAADGL	PTIEGKELSPSNCMSLLAVLPAA	TNNKGS 150
Db	85	VLGRGHSILTPENIRTSITNGAALSNSAADGL	PTIEGKELSPSNCMSLLAVLPAA	TNNKGS 144
Qy	151	QTPTTSTPNSGTTYSKTDLLLLLNNEKPSF	SNLVSGDGGIDA	DKSLTVQGISKL
Db	145	QTPTTSTPNSGTTYSKTDLLLLLNNEKPSF	SNLVSGDGGIDA	DKSLTVQGISKL
Qy	211	NTAQDGGACQCVTSPS	AMANEAPIAFVANVAGVRGGGIAA	VQDQCGQVSSSTSTEDPW 270
Db	205	NTAQDGGACQCVTSPS	AMANEAPIAFVANVAGVRGGGIAA	VQDQCGQVSSSTSTEDPW 264
Qy	271	SPSRNTAVEFDGNVARVGGGIYSYGNVAFL	NNGKTLPLNNVASPYVIAAKQ	PTSQASNT 330

Db	265	SFSRNTAV	FEFDGNVAR	VGGGIYSY	YGNVAF	LNNKGKTL	FLNNV	ASVPYIA	IAAKQ	PTSGQ	ASNT	324																																													
Qy	331	SNNTY	GDGGAIF	CKNGAQ	AGSNNS	GSVS	FDG	GVVFP	PFSSN	AAAGK	GGAIA	YAKKL	SVANC	CP	390																																										
Db	325	SNNTY	GDGGAIF	CKNGAQ	AGSNNS	GSVS	FDG	GVVFP	PFSSN	AAAGK	GGAIA	YAKKL	SVANC	CP	384																																										
Qy	391	VQFLR	NIAND	GGAII	YLGES	GELS	LSADY	GDI	IFD	GNL	KRTAK	ENAA	DVNG	VTV	SSQAISM	450																																									
Db	385	VQFLR	NIAND	GGAII	YLGES	GELS	LSADY	GDI	IFD	GNL	KRTAK	ENAA	DVNG	VTV	SSQAISM	444																																									
Qy	451	GSGGK	ITTL	RAK	AHQ	IL	FND	PI	EMANG	NN	PAQ	SSK	LLK	IND	GGY	GT	DI	IV	FANG	SSTL	510																																				
Db	445	GSGGK	ITTL	RAK	AHQ	IL	FND	PI	EMANG	NN	PAQ	SSK	LLK	IND	GGY	GT	DI	IV	FANG	SSTL	504																																				
Qy	511	YQNTI	EOGR	IV	LRE	AK	KL	VNS	LS	Q	TG	SS	LYME	AGS	TW	D	F	VP	PO	QPP	PAAN	QLIT	LS	570																																	
Db	505	YQNTI	EOGR	IV	LRE	AK	KL	VNS	LS	Q	TG	SS	LYME	AGS	TL	D	F	VP	PO	QPP	PAAN	QLIT	LS	564																																	
Qy	571	NLHL	SS	LL	ANNA	V	T	P	T	P	P	PAQ	D	SH	PA	V	I	G	S	T	T	A	G	S	V	I	T	S	G	I	F	F	E	D	L	D	T	A	Y	D	R	Y	630														
Db	565	NLHL	SS	LL	ANNA	V	T	P	T	P	P	PAQ	D	SH	PA	V	I	G	S	T	T	A	G	S	V	I	T	S	G	I	F	F	E	D	L	D	T	A	Y	D	R	Y	624														
Qy	631	WLG	NQ	K	I	N	V	L	K	L	O	G	T	K	P	P	A	N	A	P	S	D	L	T	G	N	E	M	P	K	Y	Q	G	S	N	K	L	A	W	D	N	T	A	N	N	G	P	T	L	K	A	690					
Db	625	WLG	NQ	K	I	N	V	L	K	L	O	G	T	K	P	P	A	N	A	P	S	D	L	T	G	N	E	M	P	K	Y	Q	G	S	N	K	L	A	W	D	N	T	A	N	N	G	P	T	L	K	A	684					
Qy	691	TW	T	K	T	G	N	P	G	P	E	R	V	A	S	I	V	P	N	S	L	M	G	S	I	L	D	I	R	S	A	H	S	A	I	Q	A	S	V	D	G	R	S	C	R	G	L	M	V	G	S	V	N	F	F	H	750
Db	685	TW	T	K	T	G	N	P	G	P	E	R	V	A	S	I	V	P	N	S	L	M	G	S	I	L	D	I	R	S	A	H	S	A	I	Q	A	S	V	D	G	R	S	C	R	G	L	M	V	G	S	V	N	F	F	H	744
Qy	751	DR	D	A	L	G	Q	G	Y	R	I	S	G	G	S	I	G	A	N	S	Y	P	G	S	N	F	L	A	F	T	E	V	F	O	R	S	K	D	Y	V	V	C	R	N	H	H	A	C	I	G	S	V	Y	L	S	810	
Db	745	DR	D	A	L	G	Q	G	Y	R	I	S	G	G	S	I	G	A	N	S	Y	P	G	S	N	F	L	A	F	T	E	V	F	O	R	S	K	D	Y	V	V	C	R	N	H	H	A	C	I	G	S	V	Y	L	S	804	
Qy	811	TQ	A	L	C	G	S	Y	L	F	D	A	P	I	R	A	S	Y	F	G	F	G	N	H	M	K	T	S	T	T	A	B	E	S	D	V	R	D	N	N	C	L	A	G	E	I	G	A	G	L	P	I	V	I	T	P	870
Db	805	TQ	A	L	C	G	S	Y	L	F	D	A	P	I	R	A	S	Y	F	G	F	G	N	H	M	K	T	S	T	T	A	B	E	S	D	V	R	D	N	N	C	L	A	G	E	I	G	A	G	L	P	I	V	I</			

```

RESULT 9
US-09-620-412C-190
; Sequence 190, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

```

Query Match	96.6%	Score 5090;	DB 2;	Length 1006;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 979; Conservative	2;	Mismatches	1;	Indels 0; Gaps 0

```
QY 31 MYPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTKNLDNSIAALPLSCFNLLGSFT 90
Db 25 MIPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTKNLDNSIAALPLSCFNLLGSFT 84
QY 91 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 150
Db 85 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 144
QY 151 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 210
Db 145 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 204
QY 211 NTAQADGGACQVVTFSANAEAPAFVANVAVRGGGIAAVQDGOOGVSSSTSTEDPV 270
Db 205 NTAQADGGACQVVTFSANAEAPAFVANVAVRGGGIAAVQDGOOGVSSSTSTEDPV 264
QY 271 SFSRNTAVEFDGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQASNT 330
Db 265 SFSRNTAVEFDGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQASNT 324
QY 331 SNNGYDGGAI FCKNGAQAGNNNSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVANCGP 390
Db 325 SNNGYDGGAI FCKNGAQAGNNNSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVANCGP 384
QY 391 VOFLRNANDGGAIIYLGESGELSADYGDIIIPDGNLKTAKENADVNGVTVSSQAISM 450
Db 385 VOFLRNANDGGAIIYLGESGELSADYGDIIIPDGNLKTAKENADVNGVTVSSQAISM 444
QY 451 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
Db 445 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 504
QY 511 YQNVTTIEQRI VLRREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPQPPAANQLITLS 570
Db 505 YQNVTTIEQRI VLRREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPQPPAANQLITLS 564
QY 571 NLHLSLSSLLANNAVTPNPTNPAQDSHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDRYD 630
Db 565 NLHLSLSSLLANNAVTPNPTNPAQDSHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDRYD 624
QY 631 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 690
Db 625 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 684
QY 691 TWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 750
Db 685 TWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 744
QY 751 YSFMAAYICDAYRTISGTETTLTSHQETWTTDAPHLARHGVRGSMYASLTNSIEVYGH 984
Db 991 GRYEYRDASRGYGLSAGSRVRF 1012
Db 985 GRYEYRDASRGYGLSAGSRVRF 1006
```

RESULT 10

US-09-598-419-190

; Sequence 190, Application US/09598419

```
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-190
```

```
Query Match 96.6%; Score 5090; DB 2; Length 1006;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 31 MYPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTKNLDNSIAALPLSCFNLLGSFT 90
Db 25 MIPOGIYDGETLVSPFYTVIGDPSTTVFSAGELTKNLDNSIAALPLSCFNLLGSFT 84
QY 91 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 150
Db 85 VLGRGHSHTFENIRSTNGAALSNSAADGLFTIEGKELSFNCNSLLAVLPAATTNKG 144
QY 151 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 210
Db 145 QPPTTTSPNSGTIYKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 204
QY 211 NTAQADGGACQVVTFSANAEAPAFVANVAVRGGGIAAVQDGOOGVSSSTSTEDPV 270
Db 205 NTAQADGGACQVVTFSANAEAPAFVANVAVRGGGIAAVQDGOOGVSSSTSTEDPV 264
QY 271 SFSRNTAVEFDGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQASNT 330
Db 265 SFSRNTAVEFDGNVARVGGIYSGNVAFNNKGTFLNNVASPVVIAAKQPTSGQASNT 324
QY 331 SNNGYDGGAI FCKNGAQAGNNNSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVANCGP 390
Db 325 SNNGYDGGAI FCKNGAQAGNNNSGVSFDPGEGVVPFSSNVAAGKGAIIYAKKLSVANCGP 384
QY 391 VOFLRNANDGGAIIYLGESGELSADYGDIIIPDGNLKTAKENADVNGVTVSSQAISM 450
Db 385 VOFLRNANDGGAIIYLGESGELSADYGDIIIPDGNLKTAKENADVNGVTVSSQAISM 444
QY 451 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
Db 445 GSGGKITTURAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 504
QY 511 YQNVTTIEQRI VLRREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPQPPAANQLITLS 570
Db 505 YQNVTTIEQRI VLRREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPQPPAANQLITLS 564
QY 571 NLHLSLSSLLANNAVTPNPTNPAQDSHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDRYD 630
Db 565 NLHLSLSSLLANNAVTPNPTNPAQDSHPAVIGSTTAGSVTTISGPIFFEDLDDTAYDRYD 624
QY 631 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 690
Db 625 WLGSNOKINVLKQLGTCKPPANAPSDTLGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 684
QY 691 TWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 750
Db 685 TWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 744
QY 751 DRDALGGYRYISGGYSLGANSYFGSSMFLAFTVEFGRSKDYVVCRSNHHACIGSYLS 810
Db 745 DRDALGGYRYISGGYSLGANSYFGSSMFLAFTVEFGRSKDYVVCRSNHHACIGSYLS 804
```

QY 811 TQALCGSYLFGDAPFIRASYFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 870
DB 805 TQALCGSYLFGDAPFIRASYFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 864
QY 871 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLNLSVPVGVKFCRCSSTHPNK 930
DB 865 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLNLSVPVGVKFCRCSSTHPNK 924
QY 931 YSFMAAYICDAYRTISGTETLLSHOETWTTDAPHLARHGVVVRGSMYASLTSNIEVYGH 990
DB 925 YSFMAAYICDAYRTISGTETLLSHOETWTTDAPHLARHGVVVRGSMYASLTSNIEVYGH 984
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
DB 985 GRYEYRDASRGYGLSAGSKVRF 1006
RESULT 11
US-09-556-877-176
; Sequence 176, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-556-877-176

Query Match 96.5%; Score 5084; DB 2; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 31 MVPOGIYDGETLTVSPFPYTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 90
DB 1 MIPOGIYDGETLTVSPFPYTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 60
QY 91 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGFKELSFSCNCSNLLAVLPAAATNKGK 150
DB 61 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGFKELSFSCNCSNLLAVLPAAATNKGK 120
QY 151 QTPTTTSPSNGTIYSKTDLLLNNEKPSFYSNLVSGDGAIDAKSLTVQGISKLCVPOE 210
DB 121 QTPTTTSPSNGTIYSKTDLLLNNEKPSFYSNLVSGDGAIDAKSLTVQGISKLCVPOE 180
QY 211 NTAQDGGACQVWTSFSMAHEAPAFVANVAGVRGGIAAVQDQCGVSSSTSTEDPVV 270
DB 181 NTAQDGGACQVWTSFSMAHEAPAFVANVAGVRGGIAAVQDQCGVSSSTSTEDPVV 240
QY 271 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 330
DB 241 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 300
QY 331 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFFSSNVAAGKGAIYAKKLSVANCGP 390
DB 301 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFFSSNVAAGKGAIYAKKLSVANCGP 360
QY 391 VOPLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAISM 450

DB 361 VOPLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAISM 420
QY 451 GSGGKIITTLRAKAGHQLFNDPIEMANGNPOAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
DB 421 GSGGKIITTLRAKAGHQLFNDPIEMANGNPOAQSSKLLKINDGEGYTGDIIVFANGSSTL 480
QY 511 YQNTYISQGRIVLREKAKLSVNSISQSGSYMEAGSTMDPVTPOPPQPPAANQLITLS 570
DB 481 YQNTYISQGRIVLREKAKLSVNSISQSGSYMEAGSTMDPVTPOPPQPPAANQLITLS 540
QY 571 NLHLSLSLLANNAVTPPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD 630
DB 541 NLHLSLSLLANNAVTPPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD 600
QY 631 WLGSNQKINVKLQGTGKPPANAPSDTLTGNEMPKYGYQGSWKLAWDPTANNPGYTLKA 690
DB 601 WLGSNQKINVKLQGTGKPPANAPSDTLTGNEMPKYGYQGSWKLAWDPTANNPGYTLKA 660
QY 691 TWTGTGYNRGPVVASLVPNSLWGSIIIDIRSAHSAIQASVDGRSYCRGLWVSGVSNRPTH 750
DB 661 TWTGTGYNRGPVVASLVPNSLWGSIIIDIRSAHSAIQASVDGRSYCRGLWVSGVSNRPTH 720
QY 751 DRDALGQGYRYISGYSILGANSYFGSSMPGLAPTEVFGRSKDYVVCRSNHHACIGSVYLS 810
DB 721 DRDALGQGYRYISGYSILGANSYFGSSMPGLAPTEVFGRSKDYVVCRSNHHACIGSVYLS 780
QY 811 TQALCGSYLFGDAPFIRASYFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 870
DB 781 TQALCGSYLFGDAPFIRASYFGNOMKTSYTFABESDVRWNNCLAGEIGAGLPVITP 840
QY 871 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLNLSVPVGVKFCRCSSTHPNK 930
DB 841 SKLYLNELRPFVQAFBSYADHESFTTEGDQARAFKSGHLNLSVPVGVKFCRCSSTHPNK 900
QY 931 YSFMAAYICDAYRTISGTETLLSHOETWTTDAPHLARHGVVVRGSMYASLTSNIEVYGH 990
DB 901 YSFMAAYICDAYRTISGTETLLSHOETWTTDAPHLARHGVVVRGSMYASLTSNIEVYGH 960
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
DB 961 GRYEYRDASRGYGLSAGSKVRF 982
RESULT 12
US-09-620-412C-176
; Sequence 176, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-412C-176
Query Match 96.5%; Score 5084; DB 2; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 31 MVPOGIYDGETLTVSPFPYTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 90
DB 1 MIPOGIYDGETLTVSPFPYTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 60
QY 91 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGFKELSFSCNCSNLLAVLPAAATNKGK 150
DB 61 VLGRGHSITFENIRTSNGAALSNSAADGLFTIEGFKELSFSCNCSNLLAVLPAAATNKGK 120
QY 151 QTPTTTSPSNGTIYSKTDLLLNNEKPSFYSNLVSGDGAIDAKSLTVQGISKLCVPOE 210
DB 121 QTPTTTSPSNGTIYSKTDLLLNNEKPSFYSNLVSGDGAIDAKSLTVQGISKLCVPOE 180
QY 211 NTAQDGGACQVWTSFSMAHEAPAFVANVAGVRGGIAAVQDQCGVSSSTSTEDPVV 270
DB 181 NTAQDGGACQVWTSFSMAHEAPAFVANVAGVRGGIAAVQDQCGVSSSTSTEDPVV 240
QY 271 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 330
DB 241 SFSRNTAVEFDGNARVGGIYSYGNVAPLNNKTLFLNNVASPVYIAAKOPTSQASNT 300
QY 331 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFFSSNVAAGKGAIYAKKLSVANCGP 390
DB 301 SNNTYDGGAIIFCKGQAQAGSNNSGVSFDGEGVFFSSNVAAGKGAIYAKKLSVANCGP 360
QY 391 VOPLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENADVNGVTVSSQAISM 450

Db 1 MIPQIYDGETLTVSPYTVIGDPGCTTVFSGAGELTKNLDNSIAALPLSCFNGLLGSFT 60
Qy 91 VLGRGHSITPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 150
Db 61 VLGRGHSITPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 120
Qy 151 QTPPTTSTPSNGTIYSKTDLLLLNNEKFSFYNSLVSGDGAIDAKSLTVQGISKLCVQOE 210
Db 121 QTPPTTSTPSNGTIYSKTDLLLLNNEKFSFYNSLVSGDGAIDAKSLTVQGISKLCVQOE 180
Qy 211 NTAQADGGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 270
Db 181 NTAQADGGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 240
Qy 271 SPSRNTAVEFDGNVARGVGGIYSYGNVAFVNNKTLFLNNVASPVYIAAKQPTSGOANT 330
Db 241 SPSRNTAVEFDGNVARGVGGIYSYGNVAFVNNKTLFLNNVASPVYIAAKQPTSGOANT 300
Qy 331 SNNGYDGGAI FCKNGAQAGSNNSGVSFDPGEGVVFSSNVAAGKGAIVAKKLSVANCGP 390
Db 301 SNNGYDGGAI FCKNGAQAGSNNSGVSFDPGEGVVFSSNVAAGKGAIVAKKLSVANCGP 360
Qy 391 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVNGVTVSSQAISM 450
Db 361 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVNGVTVSSQAISM 420
Qy 451 GSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
Db 421 GSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 480
Qy 511 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANOLITLS 570
Db 481 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANOLITLS 540
Qy 571 NLHLSLSLLANNAVTPNPPAODSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYRDY 630
Db 541 NLHLSLSLLANNAVTPNPPAODSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYRDY 600
Qy 631 WLGSNOKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 690
Db 601 WLGSNOKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 660
Qy 691 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 750
Db 661 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 720
Qy 751 DRDALGQGYRIYISGGYSLGANSYFGSMFGLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 810
Db 721 DRDALGQGYRIYISGGYSLGANSYFGSMFGLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 780
Qy 811 TQOALCGSYLFGDAFIRASYFGNQHMKTSYTPAESDVRWNNCLAGEIGAGLPVITP 870
Db 781 TQOALCGSYLFGDAFIRASYFGNQHMKTSYTPAESDVRWNNCLAGEIGAGLPVITP 840
Qy 871 SKLYNELRPVQAEFSYADHESFTEGDOARAPKSHLNLSPVGVKFDRCSSHTPNK 930
Db 841 SKLYNELRPVQAEFSYADHESFTEGDOARAPKSHLNLSPVGVKFDRCSSHTPNK 900
Qy 931 YSPMAAIVCDAYRTISGTETTLTSHOBTWTTDAPHLARHGCVVVRGSMYASLTNSIEVYGH 990
Db 901 YSPMAAIVCDAYRTISGTETTLTSHOBTWTTDAPHLARHGCVVVRGSMYASLTNSIEVYGH 960
Qy 991 GRYEYRDASRGYGLSAGSRVRF 1012
Db 961 GRYEYRDASRGYGLSAGSKVXF 982

RESULT 13

US-09-598-419-176

; Sequence 176, Application US/09598419

; Patent No. 6565856

; GENERAL INFORMATION:

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-598-419-176

Query Match 96.5%; Score 5084; DB 2; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 31 MIPQIYDGETLTVSPYTVIGDPGCTTVFSGAGELTKNLDNSIAALPLSCFNGLLGSFT 90
Db 1 MIPQIYDGETLTVSPYTVIGDPGCTTVFSGAGELTKNLDNSIAALPLSCFNGLLGSFT 60
Qy 91 VLGRGHSITPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 150
Db 61 VLGRGHSITPENIRISTNGAALSNSAADGLFTIEGFKELSPNSCNLSLAVLPAATTNKS 120
Qy 151 QTPPTTSTPSNGTIYSKTDLLLLNNEKFSFYNSLVSGDGAIDAKSLTVQGISKLCVQOE 210
Db 121 QTPPTTSTPSNGTIYSKTDLLLLNNEKFSFYNSLVSGDGAIDAKSLTVQGISKLCVQOE 180
Qy 211 NTAQADGGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 270
Db 181 NTAQADGGACQVVTFSFAMANEAPAFVANVAGVGGGIAAVODGQGVSSSTSTEDPVV 240
Qy 271 SPSRNTAVEFDGNVARGVGGIYSYGNVAFVNNKTLFLNNVASPVYIAAKQPTSGOANT 330
Db 241 SPSRNTAVEFDGNVARGVGGIYSYGNVAFVNNKTLFLNNVASPVYIAAKQPTSGOANT 300
Qy 331 SNNGYDGGAI FCKNGAQAGSNNSGVSFDPGEGVVFSSNVAAGKGAIVAKKLSVANCGP 390
Db 301 SNNGYDGGAI FCKNGAQAGSNNSGVSFDPGEGVVFSSNVAAGKGAIVAKKLSVANCGP 360
Qy 391 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVNGVTVSSQAISM 450
Db 361 VQFLRNANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVNGVTVSSQAISM 420
Qy 451 GSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
Db 421 GSGGKITTLRAKAGHOILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 480
Qy 511 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANOLITLS 570
Db 481 YQNVITIEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANOLITLS 540
Qy 571 NLHLSLSLLANNAVTPNPPAODSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYRDY 630
Db 541 NLHLSLSLLANNAVTPNPPAODSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYRDY 600
Qy 631 WLGSNOKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 690
Db 601 WLGSNOKINVLKQLGKTPPANAPSDLTLGNEMPKYGYOGSKWLAWDPNTANNPPTLKA 660
Qy 691 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 750
Db 661 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 720
Qy 751 DRDALGQGYRIYISGGYSLGANSYFGSMFGLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 810
Db 721 DRDALGQGYRIYISGGYSLGANSYFGSMFGLAFTEVFGRSKDYVVCRSNHHACIGSVYLS 780

QY 811 TQOALCGSYLFGDAFIRASYGFGNOHMTSYTFAESDVRDNNCLAGEIGAGLPIVITP 870
Db 781 TQOALCGSYLFGDAFIRASYGFGNOHMTSYTFAESDVRDNNCLAGEIGAGLPIVITP 840
QY 871 SKLYLNELRPPVQAEFSYADHESFTEBGDQARAFKSGHLLNLSVPGVKFDRCSSTHFNK 930
Db 841 SKLYLNELRPPVQAEFSYADHESFTEBGDQARAFKSGHLLNLSVPGVKFDRCSSTHFNK 900
QY 931 YSFMAAYICDAVRTISGTETTLTLLSHQETWTTDAPFLARHGVVVRGSMYASLTNSIEVYGH 990
Db 901 YSFMAAYICDAVRTISGTETTLTLLSHQETWTTDAPFLARHGVVVRGSMYASLTNSIEVYGH 960
QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
Db 961 GRYEYRDASRGYGLSAGSKVXF 982

RESULT 14

US-10-197-220-169

; Sequence 169, Application US/10197220

; Patent No. 6919187

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Guderian, Jeff

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Maisonneuve, Jean-Francois L.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT

; FILE REFERENCE: 210121.515C3

; CURRENT APPLICATION NUMBER: US/10/197,220

; CURRENT FILING DATE: 2002-07-15

; NUMBER OF SEQ ID NOS: 175

; SEQ ID NO 169

; LENGTH: 670

; TYPE: PRT

; ORGANISM: Chlamydia

US-10-197-220-169

Query Match 63.3%; Score 3336.5; DB 2; Length 670;

Best Local Similarity 97.2%; Pred. No. 3.1e-283;

Matches 651; Conservative 11; Mismatches 7; Indels 1; Gaps 1;

QY 28 AEIMVPOGIYDGETLTVSFPYTVIGDPSTGVFSAGELTLKNLONSIAALPLSCFNLG 87
Db 1 AEIMVPOGIYDGETLTVSFPYTVIGDPSTGVFSAGELTLKNLONSIAALPLSCFNLG 60
QY 88 SFTVLGRGHSITFENIRSTNGAALSNSAADGLFTIEGKELSFNCNLSLLAVLPAATTN 147
Db 61 SFTVLGRGHSITFENIRSTNGAALSNSAADGLFTIEGKELSFNCNLSLLAVLPAATTN 120
QY 148 KGSQPTTTTSPNGTIYSKTDLLLNNEKSFYSNLVSGDGAIDAKSLTVQGISKLCV 207
Db 121 KGSQPTTTTSPNGTIYSKTDLLLNNEKSFYSNLVSGDGAIDAKSLTVQGISKLCV 180
QY 208 FOENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGIAAVQDQOQGVSSSTSTED 267
Db 181 FOENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGIAAVQDQOQGVSSSTSTED 240
QY 268 VVPSFRNTAVFDGNVARVGGIYSGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQA 327
Db 241 VVPSFRNTAVFDGNVARVGGIYSGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQA 300
QY 328 SNTSNYGDGGAIFCKNGAQ-AGSNNSGVSFDGEGVVFFSNNVAGKGGAIYAKKLSVA 386
Db 301 SNTSNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFFSNNVAGKGGAIYAKKLSVA 360
QY 387 NCGPVQFLNNTANDGGAIIYLGESGELSISADYGDIIIFDGNLKRTAKENAADVNGVTSSQ 446
Db 361 NCGPVQFLNNTANDGGAIIYLGESGELSISADYGDIIIFDGNLKRTAKENAADVNGVTSSQ 420
QY 447 AISMGSGGKITTIRAKAGHOILFNDPIEMANGNNOQAOSSEPLKINDGEGYTGDIVPANG 506

Db 421 AISMGSGGKITTIRAKAGHOILFNDPIEMANGNNOQAOSSEPLKINDGEGYTGDIVPANG 480
QY 507 SSTLYQNVITIEQGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDPVTPPQPPQPPAANQL 566
Db 481 NSTLYQNVITIEQGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDPVTPPQPPQPPAANQL 540
QY 567 ITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626
Db 541 ITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 600
QY 627 DRYDMLGSNOKINVLKLGTKPPANAPSDLTILGNEMPKYQYQGSWKLAWDPNTANNPGY 686
Db 601 DRYDMLGSNOKINVLKLGTKPPANAPSDLTILGNEMPKYQYQGSWKLAWDPNTANNPGY 660
QY 687 TLKATWTKTG 696
Db 661 TLKATWTKTG 670

RESULT 15

US-09-612-402B-17

; Sequence 17, Application US/09612402B

; Patent No. 6642023

; GENERAL INFORMATION:

; APPLICANT: Jackson, W. James

; APPLICANT: Pace, John

; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

; Patent No. 6642023

; FILE REFERENCE: 7969-086-999

; CURRENT APPLICATION NUMBER: US/09/612,402B

; CURRENT FILING DATE: 2000-07-06

; PRIOR APPLICATION NUMBER: 08/942,596

; PRIOR FILING DATE: 1997-10-02

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 17

; LENGTH: 505

; TYPE: PRT

; ORGANISM: Chlamydia sp.

US-09-612-402B-17

Query Match 48.5%; Score 2552; DB 2; Length 505;

Best Local Similarity 100.0%; Pred. No. 1.1e-214;

Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPOGIYDGETLTVSFPYTVIGDPSTGVFSAGELTLKNLONSIAALPLSCFNLG 88
Db 1 EIMVPOGIYDGETLTVSFPYTVIGDPSTGVFSAGELTLKNLONSIAALPLSCFNLG 60
QY 89 FTVLGRGHSITFENIRSTNGAALSNSAADGLFTIEGKELSFNCNLSLLAVLPAATTN 148
Db 61 FTVLGRGHSITFENIRSTNGAALSNSAADGLFTIEGKELSFNCNLSLLAVLPAATTN 120
QY 149 GSOPTTTTSPNGTIYSKTDLLLNNEKSFYSNLVSGDGAIDAKSLTVQGISKLCV 208
Db 121 GSOPTTTTSPNGTIYSKTDLLLNNEKSFYSNLVSGDGAIDAKSLTVQGISKLCV 180
QY 209 QENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGIAAVQDQOQGVSSSTSTEDP 268
Db 181 QENTAQADGGACQVVTFSAMANEAPAFVANVAGVRGGIAAVQDQOQGVSSSTSTEDP 240
QY 269 VVPSFRNTAVFDGNVARVGGIYSGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQA 328
Db 241 VVPSFRNTAVFDGNVARVGGIYSGNVAFPLNNGKTLFLNNVASPVYIAAKOPTSQA 300
QY 329 NTSNNGYDGGAI FCKNGAQAGSNNSGVSFDGEGVVFFSNNVAGKGGAIYAKKLSVANC 388
Db 301 NTSNNGYDGGAI FCKNGAQAGSNNSGVSFDGEGVVFFSNNVAGKGGAIYAKKLSVANC 360
QY 389 GPVQFLNNTANDGGAIIYLGESGELSISADYGDIIIFDGNLKRTAKENAADVNGVTSSQAI 448
Db 361 GPVQFLNNTANDGGAIIYLGESGELSISADYGDIIIFDGNLKRTAKENAADVNGVTSSQAI 420

QY 449 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 508
DB 421 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 480
QY 509 TLYQNVTTIEQGRIVLRKAKLSVNS 533
DB 481 TLYQNVTTIEQGRIVLRKAKLSVNS 505

RESULT 16
US-09-542-520-17
; Sequence 17, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John L.
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-542-520-17

Query Match 48.4%; Score 2547; DB 2; Length 505;
Best Local Similarity 99.8%; Pred. No. 3.1e-214; Indels 0; Gaps 0;
Matches 504; Conservative 1; Mismatches 0;

QY 29 EIMVPGIYDGETLTVSPFTVIGDPSGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 88
DB 1 EIMVPGIYDGETLTVSPFTVIGDPSGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 60

QY 89 FTVLGRHSLTFENIRSTNGAALSADGLFTTBGFKLSFNSCNLSLAVLPAATTNK 148
DB 61 FTVLGRHSLTFENIRSTNGAALSADGLFTTBGFKLSFNSCNLSLAVLPAATTNK 120

QY 149 GSOTPTTSTPSNGTIIYSKTDLLANNEKESFYSLVSGDGAIDAKSLTVQGISKLCVF 208
DB 121 GSOTPTTSTPSNGTIIYSKTDLLANNEKESFYSLVSGDGAIDAKSLTVQGISKLCVF 180

QY 209 QENTAQADGACQVWTSFSAMANEAPIAFVANVAGVRGGIAAVQDQGGVSSSTSTEDP 268
DB 181 QENTAQADGACQVWTSFSAMANEAPIAFVANVAGVRGGIAAVQDQGGVSSSTSTEDP 240

QY 269 VVSFSRNTAYEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQAS 328
DB 241 VVSFSRNTAYEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQAS 300

QY 329 NTSNNYGDGAIKFCNGAQAGSNNSGVSFDGEGVFPSSNVAAGKGAIVAKLSVANC 388
DB 301 NTSNNYGDGAIKFCNGAQAGSNNSGVSFDGEGVFPSSNVAAGKGAIVAKLSVANC 360

QY 389 GPVQFLNIANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI 448
DB 361 GPVQFLNIANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

QY 449 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 508
DB 421 SMGSGKITTIRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIVFANGSS 480

QY 509 TLYQNVTTIEQGRIVLRKAKLSVNS 533
DB 481 TLYQNVTTIEQGRIVLRKAKLSVNS 505

RESULT 17
US-09-612-402B-36

; Sequence 36, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; Patent No. 6642023
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-36

Query Match 44.6%; Score 2350; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 5e-197;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GCACQVWTSFSAMANEAPIAFVANVAGVRGGIAAVQDQGGVSSSTSTEDPVSFSRNT 276
DB 1 GCACQVWTSFSAMANEAPIAFVANVAGVRGGIAAVQDQGGVSSSTSTEDPVSFSRNT 60

QY 277 AVEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 336
DB 61 AVEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 120

QY 337 GCAIFCKNGAQAGSNNSGVSFDGEGVFPSSNVAAGKGAIVAKLSVANGCPVQFLRN 396
DB 121 GCAIFCKNGAQAGSNNSGVSFDGEGVFPSSNVAAGKGAIVAKLSVANGCPVQFLRN 180

QY 397 IANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI SMGSGKI 456
DB 181 IANDGAIYLGSGELSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI SMGSGKI 240

QY 457 TTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIYPANGSSTLYQNVTI 516
DB 241 TTLRAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIYPANGSSTLYQNVTI 300

QY 517 EGRIVLRKAKLSVNSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI SMGSGKI 576
DB 301 EGRIVLRKAKLSVNSLSDYDGIIFDGNLKRKTAKENAADVNGVTVSSQAI SMGSGKI 360

QY 577 SLLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQ 636
DB 361 SLLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQ 420

QY 637 KINVLKQLGTGKPPANAPSDLTILGNEMPKYQGSWKL 674
DB 421 KINVLKQLGTGKPPANAPSDLTILGNEMPKYQGSWKL 458

RESULT 18
US-09-542-520-36
; Sequence 36, Application US/09542520
; Patent No. 6887843
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John L.
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-076-999
; CURRENT APPLICATION NUMBER: US/09/542,520
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: PCT/US98/20737
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36

```
/ LENGTH: 458
/ TYPE: PRT
/ ORGANISM: Chlamydia
US-09-542-520-36

Query Match      44.6%; Score 2350; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 5e-197;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GGACQVTSFSAWANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNT 276
Db 1 GGACQVTSFSAWANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNT 60

QY 277 AVEFGNVARCGGYSYGNVAFNNKTLFLNNVASPVYIAAKOPTSGOASNTSNNGD 336
Db 61 AVEFGNVARCGGYSYGNVAFNNKTLFLNNVASPVYIAAKOPTSGOASNTSNNGD 120

QY 337 GGAIFCKGAQAGSNNSVSPDGGVVPFSSNVAAGKGAIYAKKLSVANGCPVQPLRN 396
Db 121 GGAIFCKGAQAGSNNSVSPDGGVVPFSSNVAAGKGAIYAKKLSVANGCPVQPLRN 180

QY 397 TANDGGAIIYLGSGELSLADYDGIIPDGNLKRKTAKENAADVNGVTVSSQAISMGGGKI 456
Db 181 TANDGGAIIYLGSGELSLADYDGIIPDGNLKRKTAKENAADVNGVTVSSQAISMGGGKI 240

QY 457 TTRAKAGHOILFNDPIEMANGNPOAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNTI 516
Db 241 TTRAKAGHOILFNDPIEMANGNPOAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNTI 300

QY 517 EGRIVLREKAKLSVNSLSQTCGSLYMEAGSTWDFVTPPOQPPAANQLITLSNLHLSL 576
Db 301 EGRIVLREKAKLSVNSLSQTCGSLYMEAGSTWDFVTPPOQPPAANQLITLSNLHLSL 360

QY 577 SELLANNAVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFPEDLDDTAYDRYDMLGSNQ 636
Db 361 SELLANNAVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFPEDLDDTAYDRYDMLGSNQ 420

QY 637 KINVLKLOLGTGPPANAPSDLTGNEPKYGYQGSWKL 674
Db 421 KINVLKLOLGTGPPANAPSDLTGNEPKYGYQGSWKL 458

RESULT 19
US-09-612-402B-37
; Sequence 37, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Pace, John
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-37

Query Match      32.9%; Score 1735; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.3e-143;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNF 747
Db 1 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNF 60

QY 748 FYHRRDALGQGYRIISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHHACIGSV 807
Db 61 FYHRRDALGQGYRIISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHHACIGSV 120

QY 808 YLSTQOALCGSYLFGDAFIRASYGFGNQHKMTSYTFAEESDVRDNNCLAGEICAGLPV 867
Db 121 YLSTQOALCGSYLFGDAFIRASYGFGNQHKMTSYTFAEESDVRDNNCLAGEICAGLPV 180

QY 868 ITPSKLYLNELRPFVQAEPFSAHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 927
Db 181 ITPSKLYLNELRPFVQAEPFSAHESFTEBGDQARAFKSGHLLNLSVPVGVKPFDRCSSTH 240

QY 928 PNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAFHARHGTVVGRGSMYASLTSTNIEV 987
Db 241 PNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAFHARHGTVVGRGSMYASLTSTNIEV 300

QY 988 YGHRGYEYRDASRGYGLSAGSRVRF 1012
Db 301 YGHRGYEYRDASRGYGLSAGSRVRF 325

RESULT 21
US-09-438-185A-455
; Sequence 455, Application US/09438185A

Query Match      32.9%; Score 1735; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.3e-143;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNF 747
Db 1 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNF 60

QY 748 FYHRRDALGQGYRIISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRRNHHACIGSV 807
```

```
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0453
; US-09-438-185A-455

Query Match      25.9%; Score 1365.5; DB 2; Length 999;
Best Local Similarity 33.5%; Pred. No. 3.8e-110;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY      1  MOTSFHKFPLSMILAYSCCSLNGGVAABIMVQGIYDGETLTVSPYTVIGDPGSGTVF 60
Db      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      27  MKTSIRKFLISTTLAPCFAST---AFTVEVIMSENFDDGSSGKI-FPYTILSPRGLTLCI 82
Db      27  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61  SAGELTKNLDNIAALPLSCFNLGSGFTVLGRGHSLTPENIRTSNNGAALSAAADG- 119
Db      61  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      83  FSGDLIANDLNAISRTSSCFNAGALQILKGGVFSFLNIRSSADGAAISSVITQNP 142
Db      83  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      120 ---LFTIEGKELSFNSCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNE 176
Db      120 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      143  ELCPLSFSGFSQIMFDCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 191
Db      143 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      177  KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTQAADGACQVTVSFSAMANEAPIA 236
Db      177 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      192  SILFQYNRSAGFAGAIRGTSTITTENTKKSLLFNNGSISNGALTGSAANLNNNSAPVI 251
Db      192 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      237  FVANVAGVGGIAAVQDGGQVSSSTSTEDPVVFSFRNTAVFEDGNVAVGGGIYSYGN 296
Db      237 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      252  FSTNATGIYGAIYLT-----GGSLMTS-----GNLSGVLPVNNSSRSGGAIYANGN 298
Db      252 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      297  VAPLNNKTLFLNNVASPV-YIAKQPTSGQASNTSNYDGGGAIFCNGAQAQGSNNSGS 355
Db      297 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      299  VTFENNSDLTFQNTASFQNSLPAPTPTPTTAVTP-LLGYGGAIFCTPTPTPTPTGV-S 356
Db      299 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      356  VSPDGEVGFVPSNVAAGKGAIAKLVANCGPVQFLRNANDGGAIVLBSGELSLS 415
Db      356 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      357  LTISGENSVTFLENIASEQGALYGGKISIDSNKSTIPLNGTAGKGAIALPESGELSLS 416
Db      357 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      416  ADYGDII PDGNLKTAKENADVNGVTVSSQAI SMGSGGKITTLRAKAGHQIILFNDPIEM 475
Db      416 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      417  ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKPAITLGAQTGYLYFYDPI-- 465
Db      417 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      476  ANGNQPAQSGSKLKIN-----DEGYTGDIIVF-----ANGSTLYQNTVIE 517
Db      476 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      466  TSDLSAASAATAVWNPKASADG-AYSGTIVFSGETILTATEAATPANATSTLNKOLE 524
Db      466 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      518  QGRIVLREKALSYNLSQTCGS-LYMEAGSTWDPVTPQPQPPAANQLITLSNLHLSL 576
Db      518 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      525  GGTLLARNGATLVNHNFTQDEKSVVIMDAGTT--LATNGANNTDGA---ITLKNLVNL 579
Db      525 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      577  SSLANNVATVPTNPPAQDGHPAVIG-STTAGSVTISGPPIFFEDLDITAYDRYDWLGSN 635
Db      577 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      580  DSLDGTAK-----AVVNQSTNGALTISGTLGLVKNSDCCDNHGMFKD 624
Db      580 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0453
; US-09-438-185A-455

Query Match      25.9%; Score 1365.5; DB 2; Length 999;
Best Local Similarity 33.5%; Pred. No. 3.8e-110;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY      1  MOTSFHKFPLSMILAYSCCSLNGGVAABIMVQGIYDGETLTVSPYTVIGDPGSGTVF 60
Db      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      27  MKTSIRKFLISTTLAPCFAST---AFTVEVIMSENFDDGSSGKI-FPYTILSPRGLTLCI 82
Db      27  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61  SAGELTKNLDNIAALPLSCFNLGSGFTVLGRGHSLTPENIRTSNNGAALSAAADG- 119
Db      61  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      83  FSGDLIANDLNAISRTSSCFNAGALQILKGGVFSFLNIRSSADGAAISSVITQNP 142
Db      83  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      120 ---LFTIEGKELSFNSCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNE 176
Db      120 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      143  ELCPLSFSGFSQIMFDCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 191
Db      143 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      177  KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTQAADGACQVTVSFSAMANEAPIA 236
Db      177 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      192  SILFQYNRSAGFAGAIRGTSTITTENTKKSLLFNNGSISNGALTGSAANLNNNSAPVI 251
Db      192 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      237  FVANVAGVGGIAAVQDGGQVSSSTSTEDPVVFSFRNTAVFEDGNVAVGGGIYSYGN 296
Db      237 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      252  FSTNATGIYGAIYLT-----GGSLMTS-----GNLSGVLPVNNSSRSGGAIYANGN 298
Db      252 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      297  VAPLNNKTLFLNNVASPV-YIAKQPTSGQASNTSNYDGGGAIFCNGAQAQGSNNSGS 355
Db      297 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      299  VTFENNSDLTFQNTASFQNSLPAPTPTPTTAVTP-LLGYGGAIFCTPTPTPTPTGV-S 356
Db      299 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      356  VSPDGEVGFVPSNVAAGKGAIAKLVANCGPVQFLRNANDGGAIVLBSGELSLS 415
Db      356 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      357  LTISGENSVTFLENIASEQGALYGGKISIDSNKSTIPLNGTAGKGAIALPESGELSLS 416
Db      357 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      416  ADYGDII PDGNLKTAKENADVNGVTVSSQAI SMGSGGKITTLRAKAGHQIILFNDPIEM 475
Db      416 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      417  ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKPAITLGAQTGYLYFYDPI-- 465
Db      417 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      476  ANGNQPAQSGSKLKIN-----DEGYTGDIIVF-----ANGSTLYQNTVIE 517
Db      476 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      466  TSDLSAASAATAVWNPKASADG-AYSGTIVFSGETILTATEAATPANATSTLNKOLE 524
Db      466 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      518  QGRIVLREKALSYNLSQTCGS-LYMEAGSTWDPVTPQPQPPAANQLITLSNLHLSL 576
Db      518 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      525  GGTLLARNGATLVNHNFTQDEKSVVIMDAGTT--LATNGANNTDGA---ITLKNLVNL 579
Db      525 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      577  SSLANNVATVPTNPPAQDGHPAVIG-STTAGSVTISGPPIFFEDLDITAYDRYDWLGSN 635
Db      577 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      580  DSLDGTAK-----AVVNQSTNGALTISGTLGLVKNSDCCDNHGMFKD 624
Db      580 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 455
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0453
; US-09-438-185A-455

Query Match      25.8%; Score 1359.5; DB 2; Length 973;
Best Local Similarity 33.4%; Pred. No. 1.2e-109;
Matches 349; Conservative 171; Mismatches 421; Indels 103; Gaps 27;

QY      1  MOTSFHKFPLSMILAYSCCSLNGGVAABIMVQGIYDGETLTVSPYTVIGDPGSGTVF 60
Db      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61  SAGELTKNLDNIAALPLSCFNLGSGFTVLGRGHSLTPENIRTSNNGAALSAAADG- 119
Db      61  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      57  FSGDLIANDLNAISRTSSCFNAGALQILKGGVFSFLNIRSSADGAAISSVITQNP 116
Db      57  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      120 ---LFTIEGKELSFNSCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNE 176
Db      120 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      117  ELCPLSFSGFSQIMFDCESLT-----SDTSASNVI PHASAIYATTPMLFTNND 165
Db      117 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      177  KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTQAADGACQVTVSFSAMANEAPIA 236
Db      177 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      166  SILFQYNRSAGFAGAIRGTSTITTENTKKSLLFNNGSISNGALTGSAANLNNNSAPVI 225
Db      166 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      237  FVANVAGVGGIAAVQDGGQVSSSTSTEDPVVFSFRNTAVFEDGNVAVGGGIYSYGN 296
Db      237 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

; Patent No. 6607730
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USSES THEREOF
; FILE REFERENCE: 19721-011
; CURRENT APPLICATION NUMBER: US/09/430,723
; CURRENT FILING DATE: 1999-10-29
; EARLIER APPLICATION NUMBER: 60/106,590
; EARLIER FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/133,071
; EARLIER FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-430-723-2

RESULT 22
US-09-430-723-2
; Sequence 2, Application US/09430723
; Patent No. 6607730
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USSES THEREOF
; FILE REFERENCE: 19721-011
; CURRENT APPLICATION NUMBER: US/09/430,723
; CURRENT FILING DATE: 1999-10-29
; EARLIER APPLICATION NUMBER: 60/106,590
; EARLIER FILING DATE: 1999-11-02
; EARLIER APPLICATION NUMBER: 60/133,071
; EARLIER FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 973
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-430-723-2
```

```
Db      226  FSTNATGYGGAIYIT-----GGMLTS-----GNLSGVLFVNNSRRSGGAIYANGN 272
Qy      227  VAFNNGKTLFLNNVASPV-YIAAKOPTSGQASNTSNNGYDGGAIKFCNGGAQAAGSNNSGS 355
Db      273  VTFSSNNDLTFQNTASQNSLPATPPPTPPAVTP-LLGYGGAIFCTPPATPPPTGV-S 330
Qy      356  VSPDEGVVFFSNVAAGKGGAIYAKLSVANGVQVFLRNITANDGGAIYIGSELSLS 415
Db      331  LTIGSENSVTFLENTASQGGALYKKSIDSNSKSTIFLGNITAGKGAIAIPESGELSLS 390
Qy      416  ADYGDIIFDGNLKRITAKENAAVDNGVTVSSQAISSMGSGKITTIRAKAGHOLLFNDEIEM 475
Db      391  ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKAPATLGNTOGYTIFYFDPI-- 439
Qy      476  ANGNNOQAQSSKLLKIN-----DGBYTGDIIVF-----ANGSSTLYQNVITIE 517
Db      440  TSDDLASAASAAATVVVNPVKASADG-AYSGTIVFSGETLTATBAATPATNATSTLNKLELE 498
Qy      518  QRIIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPPOPPQPPAANQLITLSNHLHL 576
Db      499  GGTALRNGATLNVHNFOTDEKSVVIMDAGTT--LATNGANNITDGA---ITLNLKLVINL 553
Qy      577  SSSLANNAVTPNPTPPAQDSHPAIG-STTAGSVTISGPIFFEDLDLDTAYDRYDMLGSN 635
Db      554  DSLDGTKA-----AVNVQSTNGALUTISGTGLVKNSQCCDNHGMFNKD 598
Qy      636  -QKINVLKLQGTKEPPANAPSL-TLGNEMPKYQGSWKLAWDNPNTANNPYTLKATWT 693
Db      599  LQQVFILSKATSNVTVTTFDSLGTNGYQOSPVGYGQGTWETIDITTT-----HTVTGNWK 653
Qy      694  KTYNPGPERVASLVPNSLWGSILDIRSAHSAIAQASVDORSY-CRGLWVSGVSNPFYHDR 752
Db      654  KTYLPHPERLAPLIPNSLWNVILDRVAVSQ--SAADGEDVPKQLSITGITNPFHANH 711
Qy      753  DALGGYRYISGYSGLANSY---FGSSMFLGAFTEVGRSKDYVVCRSNHACIGSVYL 809
Db      712  TGDARSYRHMGGYILI--NTYTRITPDAAISLGFQQLFKSKDYLVHGHNSVVPATVYS 769
Qy      810  STQQAALCG-SYLFGDFAIRASVGFQGNHMKTSYTPAESDVRWNNCLAGIBIGALPTVI 868
Db      770  NITKSLFGSSRFFSGTSRVTVSRNSKVKTSYTKLPKGRCSMNSNCLWGLEGNLPITL 829
Qy      869  TPSKLYLNLRLPFOAERSYADHESFTREGDOARAPKSGHLNLNLSVPVGVKPDRCSSHP 928
Db      830  SSRILNLKQIIPFVKAEVAYATHGGIQENTPEGRIFGHGHLINVAVPVGVFRFGKSHNR 889
Qy      929  NKYSFMAAVIDAVYTIISGTETILLSHOETWTDAPFLARHGCVVVRGSMYASLTSNIEVY 988
Db      890  DFYTIIVAYAPDVYRHNPDCTFLPINGATWTSIGNNLTRSLLVQASSHTSVNDVLEIF 949
Qy      989  GHGREYERDASRGYGLSAGSRVRF 1012
Db      950  GHGCGDIRTSEKTYLIDIGSKLRF 973
```

RESULT 23

```
US-09-428-122-2
; Sequence 2, Application US/09428122
; Patent No. 6872814
; GENERAL INFORMATION:
; APPLICANT: Connaught Laboratories Limited
; APPLICANT: Murdin et al.
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-007-019
; CURRENT APPLICATION NUMBER: US/09/428,122
; CURRENT FILING DATE: 1999-10-27
; EARLIER APPLICATION NUMBER: 60/106,046
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/132,271
; EARLIER FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 4
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-428-122-2
```

Query Match 21.4%, Score 1128; DB 2; Length 928;
Best Local Similarity 29.2%; Pred. No. 2.2e-89;
Matches 311; Conservative 169; Mismatches 394; Indels 192; Gaps 32;

```
Qy      1  MOTSFHFPFLSMILAYSCCSLNGGGYAAE-IWVPGIYDGETLTVSRPPYTVIGDPSGTVV 59
Db      1  MKSFPFPVFTFAIFPLSMI-----ATETVLDSSASFDGNK-NGNFVRSRSEDAGTY 54
Qy      60  PSAGELTKLNLDNSIAALPLSCFCGNLLGSFTVLGRGHSILTFENIRTSI-NGAALSNASAD 118
Db      55  LFKGNVTLLENIPGVTGTAITKSCFNNTKGDLTFTCGNSLLLFQTVDAQTVAAGAVNSSVVD 114
Qy      119  GLFTIEGFKELSPNCNSLLAVLPAAATNKGSOPTTTTSPNSGTIYSKTDLILLNNEKP 178
Db      115  KSTTFIGFSSLSF-----IASPGSSIITGKGAVSCSTGS-----LSLTKNVSL 157
Qy      179  SFYSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQAAGGACOVVTSFSAMANEAPAFV 238
Db      158  LFSKNFSDNGGAI-TAKTLSLTGTTMSALFSENSSKKGAIQTSDALITGNQGEVSF- 216
Qy      239  ANVAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSPSRNTAVFDCGNVARVGGIYSYGNVA 298
Db      217  -----SDNTSSDS-----GAAIFTEASVT 235
Qy      299  FLNNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAIKFCNGGAQAAGSNNSGVSF 358
Db      236  ISNNAKVSFIDN-----KVTGASSSTTGDG-SGGAICAYK-----TSTDTKVTL 278
Qy      359  DGEGVVFPSSNVAAGKGGAIYAKLSVANGCPVQVFLRNIA-----DGGAIYIGSELS 413
Db      279  TGNQMLLFSNNTTTAGGAIYVKKLELASGGLTLFSRNSVNGGTPAPKGAIAIEDSGELS 338
Qy      414  LSADYDGIIFDGNLKRITAKENAAVDNGVTVSSQAISSMGSGKITTIRAKAGHQLFLNDPI 473
Db      339  LSADSGIIVFLGN---TVTSTTPGTN-----RSSIDIGTSKMTALSAAGRAIYFYDPI 390
Qy      474  EMANGNNOQAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNVY 515
Db      391  TTGSSTT-----VTDVLKVNTPADSAQYTGNIIFTGKELSETEAASKNLTSKLLQPV 446
Qy      516  IEQGRIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPPOPPQPPAANQLITLSNHL 574
Db      447  LSGGTLSLKHGVTLTQTAQQAADSRLEMDVGTTL-----PADTS--TINNLI 494
Qy      575  SLSILLANNAVTPNPTPPAQDSHPAIGS-TTAGSVTISGPIFFEDLDLDTAYDRYDMLG 633
Db      495  NISSI-----DGAKKAKIETKATSKNLTLSGTITLLDPTGTFYENHS-LR 538
Qy      634  SNQKINVLKQL-CTKPPANAPSDLITGNEMPKYQGSW-KLAWDPNTANNPYTLKAT 691
Db      539  NPQSYDILELKASGTVTSTAVTPDPIMGEKP-HYGYQGTWGPVWGTGASTATF-----N 593
Qy      692  WTKTYNPGPERVASLVPNSLWGSILDIRSAHSAIAQASVDGRSYCRGLWVSGVSNFFVHD 751
Db      594  WTKTYIPNPERIGSLVPNSLWNAFIDISLHYLMETANEGLQDRAFCWCAGLSNFPFKD 653
Qy      752  RDALGQGVRYISGYSGLANSYFGS-SMFGLAFTFVGRSKDYVVCRSNHACIGSVYLS 810
Db      654  STKTRGFRHLSGGVYVGGNLTCTCDKLSAAFCQLFGRDREDYFVAKNQ-----GTVYGG 708
Qy      811  TQQAALCGSYLFGDAFI-----RASVGFQGNHMKTSYTFABE 846
Db      709  TLY-----YQHNETYISLPCKLRPCSLSYVPTETPVLFSGNLSYTHTDNDLTKYTYPT 763
Qy      847  SDVWDNNCLAGEIGAGLPVITPSKLYLNLRLPFOAERSYADHESFTREGDOARAPKS 906
```

Db	764	VKSGWNSDFALEFGGRAPICIDESALP-EQYMPFMKLOFVYAHQEGFKBQCTEAREFGS	822
Qy	907	GHLNLVSVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETILLSHOETWTTDAFHL	966
Db	823	SRLVNLALPIGIRFDRKESQDQATYNLTGYTVDLVRSPNPDCTTTLIRISGDSWKTGNTL	882
Qy	967	ARHGTVVRGSMVASLTSNLEVYGHGVEYRDASRGYGLSAGSRVF	1012
Db	883	ARQALVLRAGNHFCFNSNFEPASQFSFELRGSRNNDVLGAKYQF	928
RESULT 24			
US-09-198-452A-478			
; Sequence 478, Application US/09198452A			
; Patent No. 6559294			
; GENERAL INFORMATION:			
; APPLICANT: Griffiths, R.			
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection			
; FILE REFERENCE: 9710-003-999			
; CURRENT FILING DATE: 1998-11-24			
; NUMBER OF SEQ ID NOS: 6849			
; SEQ ID NO 478			
; LENGTH: 949			
; TYPE: PRT			
; ORGANISM: Chlamydia pneumoniae			
; FEATURE:			
; NAME/KEY: SITE			
; LOCATION: 1...949			
; OTHER INFORMATION: Xaa=unknown or other			
US-09-198-452A-478			
Query Match 20.7%; Score 1090; DB 2; Length 949;			
Best Local Similarity 28.8%; Pred. No. 5e-86;			
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;			
Qy	1	MQTSFHKFFLSMILAYSC--CSLNGGGYAAEIMVPGIYDGETLTVSF-----PYTVI	51
Db	22	MKTSIPMWLVSVLAFSLCHLQSLAN---EELSPDSDSFNGINDSGTFTPKTSATTYSLT	77
Qy	52	GDPSGTTVFAGELTLKLNLSIAALPL--SCFNGLLGSFTVLGRHSLFENIRISTNG	109
Db	78	GD---VFEPYRPGKT-----PLSDSCFRQTTDNLTLFLNGHSLTFGFDAGTHA	123
Qy	110	AALNSAADGLFTIEGKELSFNSCNLSLLAVLPAATTNKGSTQPTTSTPSNGTIYSKTD	169
Db	124	GAASATTTANKNLTFSGFSLISFDS-----SFTTIVTGGTILSSAGG	165
Qy	170	LLILNNEKFPYSNLVSGDGAIDAKSLTVQGISKLCVFOENTAAQADGGACQVVTSPSAM	229
Db	166	VLENIRKLVVAGNFSTADGAIGASFLLTGTSGDALFNSNSSSTKGA-----	215
Qy	230	ANEAPIAFVANVAGVRGGIAAVODGGQGVSSSTSTEDPVVPSRNT-AVEFDGNVARVG	288
Db	216	-----IATTAGAR-----IANNTGXVRFLSNIASST	241
Qy	289	GGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNGYDGGAIKCKGAQA	348
Db	242	CGAIDDEGTSILSNKKFLYP-----EGNAAKT-----GGAI-CNTKASG	280
Qy	349	G-----SNNSGVSPDGBGVVFPFNSVAAGKGAIAKLSVANCGVPQFLRNLIAND---	400
Db	281	SPELIISNKK-----TLIFASNVAEETSGGAHAKKLALSSGFTFLRNVSATP	331
Qy	401	-GGAIYLGESGELSLSADYGDIIIFDGNLKRKTAKENADVNGVTSSQAIMSGSGKITTLL	459
Db	332	KGAISIDASGELSLSAETGNITFVRNLTITT--GSTD---TPKRNAINIGSNKGKFTL	385
Qy	460	RAKAGHQILFNDPIEMANGNNQPAQSKLLKINDGEG-----YTGDIVFA-----	504
Db	386	RAAKNHTTIFYVDPI-----TSEGTSSDVLKINNGSAGALPNPYOQTILFSGETTLTADBLK	439

Db	533	-----SHTIN-----QAFTQPLVVVTAATAASDIYDALLTSPVQTPPEHYGYQGHWEAT	582
Qy	676	W-DPNTANNPPTLKATWTKTGYNCPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRS	734
Db	583	WADTSTAKSG---TMTWTTTGYNPNPERRASVVPDLSLWASFTDIRTLLQOIMTSQANSIY	638
Qy	735	YCRGLWVSGVSNPFYHDDRDLAQGVRYISGGYSLGANSY-FGSSMFLGAPTEVFGRSKDY	793
Db	639	QQRGLWASGTANPFHKDKSGTQNAFHKHSGYIVVGSADDFSENIFSVAFQQLFGKDKDL	698
Qy	794	VVCRSNHACIGSVYLSLTOQALCGSYL--FGDAF-----IRASYFGNGHMKTS	840
Db	699	FIVENTSHNYLASLYLQHRAPLGLPMPFSGITDMLKDIPLILNAQLSVSYTKNDMDTR	758
Qy	841	YTFABESDVRWNNCLAGIAGLPVITPFSKLYLNLPPVQAEFSYADHESFTEGDO	900
Db	759	YTSYPEAQGSWTNNSGALBELGSLALYLKPEAPPFQGYFPFLKFOAVYSRQONFKESGAE	818
Qy	901	ARAPKSHLLNLVSPVGVKFDRCSSTHPNKYSPMAAYICDAYRTISGTETILLSHOBTWT	960
Db	819	ARAFDDGDLVNCISIPVGIHLEKISEDEKNFEISLAYIGDVYRKNRPRSRISLWVSGASWT	878
Qy	961	TDAPFLARHGVVVRGSMVYASLTNSIEVYGHGRYERDASRGYGLSAGSRVRF	1012
Db	879	SICKNLRQAFTLASAGSHLTLSPHVELSGEAYELRGSAAHYNVDCGLRYSP	930
RESULT 27			
US-09-438-185A-448			
; Sequence 448, Application US/09438185A			
; Patent No. 6822071			
; GENERAL INFORMATION:			
; APPLICANT: Stephens, Richard			
; APPLICANT: Mitchell, Wayne			
; APPLICANT: Kalman, Sue			
; APPLICANT: Davis, Ronald			
; APPLICANT: The Regents of the University of California			
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence			
; FILE REFERENCE: 018941-000411US			
; CURRENT APPLICATION NUMBER: US/09/438,185A			
; CURRENT FILING DATE: 2002-03-13			
; PRIOR APPLICATION NUMBER: US 60/108,279			
; PRIOR FILING DATE: 1998-11-12			
; PRIOR APPLICATION NUMBER: US 60/128,606			
; PRIOR FILING DATE: 1999-04-08			
; NUMBER OF SEQ ID NOS: 1074			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 448			
; LENGTH: 938			
; TYPE: PRN			
; ORGANISM: Chlamydia pneumoniae			
; FEATURE:			
; OTHER INFORMATION: CFP0446			
US-09-438-185A-448			
Query Match 20.5%; Score 1079; DB 2; Length 938;			
Best Local Similarity 28.3%; Pred. No. 4.5e-85;			
Matches 303; Conservative 160; Mismatches 407; Indels 202; Gaps 29;			
Qy	1	MQTSFKHFFPLSMILAYSCCSLNGGYAAEI-MVPOGIVDGETLTVSPYTVIGDPGSGTV	59
Db	9	MKIPLHKLISSTLV-TPILSIATYADASLSPTDSFDGAGGSTPTPKST-ADANGTNY	66
Qy	60	FSACELTKNLDNSIAALPLSCFNGLLGSPVTLGRGSLTFENIRTSNGAALNSAAGD	119
Db	67	VLSGNVYINDAGK-TALTGCCPTTETGDLTFPGKGYSFNSPTVDAGNAGAAATYADK	125
Qy	120	LFTIEGFKELSFNSCNLSILAVLPAAITNKGSTPTTT---STPSNGTYIKTDLILLNNE	176
Db	126	ALTGTGNSLSF-----IAAPGTVASGKSTLSSAGALNLTGNTILPQNV-----	172
Qy	177	KFSFYSNLVSGDGAIDAKSLTVQGISKLQVQENTAQDAGGACQVTSFSMANEAPIA	236

RESULT 28
US-09-438-185A-447
; Sequence 447, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A

```
/ CURRENT FILING DATE: 2002-03-13
/ PRIOR APPLICATION NUMBER: US 60/108,279
/ PRIOR FILING DATE: 1998-11-12
/ PRIOR APPLICATION NUMBER: US 60/128,606
/ PRIOR FILING DATE: 1999-04-08
/ NUMBER OF SEQ ID NOS: 1074
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 447
/ LENGTH: 947
/ TYPE: PRT
/ ORGANISM: Chlamydia pneumoniae
/ FEATURE:
/ OTHER INFORMATION: Cpn0445
US-09-438-185A-447

Query Match      19.8%; Score 1041; DB 2; Length 947;
Best Local Similarity 29.0%; Pred. No. 9.9e-82;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MCTSPHKPFLSMILAYSCCSLNGGGYAAEIMV--PQGIYDGETLTVSPYTVIGDPGTT 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 MKSSVSWLFFSFIPLFSSLSI-----VAAEVLDDSSNNSYDGSNGTTFVFTTDAAGTT 67
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 59 VPSAGELTLKNDLNSIAALPLSCFQNLGSLFTVLGRGHSHTPENIRT-STNGAALSNGAA 117
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 YLLSDVDFQNGALGIPLASCFLEAGDGLTFQGNQHALKPAFINAGSSAGTAVASTAA 127
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 118 DGLFTIEGKELSFSCNLSLALVPAATNKGSOPTTTTSPNSNTIYSKTDLLLNNEK 177
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 DNKLNFDFSRLLSIISCPSLL-----SPT-----GQCALKSVDGLSLTGNSQ 170
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 178 P8PYENLVSGDGAIDAKSLTVQGISKLCVPEHTAQADGACQVVTFSAMANEAPIAF 237
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 IFTQNFSDNGGVINTKFLSGTSQFASFRNOA----- 206
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 238 VANVAGVGGGIAAQDQGGVSSSTSTEDP-VVSFSRNTAVEPDGNVARVGGGIYSGN 296
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 ---FTGKGQGVYA-----TGTITIENSPOIVFSQNLA---KGS---GGALYSTDN 249
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 297 VAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNTYGDGGAIFCGKGAQAGSNGSV 356
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 CSITDNFQVDPGNSA---WEAAQ-----AQQGAICC-----TTTDKTV 285
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 357 SPDGEVGVFFSNAAGGAIYAKLVANGCPVQPLENTAND-----GGAIYLGESG 410
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 TLTGKKNLSPTNTALTYYGGAISGLKVISAGGPTLPQSNTSGSSAGQGGGAINIASAG 345
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 411 ELSISADYGDIDFDGNLKRKTAKENAADVNGVTVSQAISMGSGGKITTLRAKAGHOILFN 470
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 ELALSATSQDITFNN-----QVTNGSTSTRNAINIDTAKVTSIRATGQSIYFY 396
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 471 DPEIANGNNOQAQSSKLLKINDRG-----YTGDIVF-----ANGSSTLYQ 512
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 397 DPIT---NPGTAASDTDLNLMLDANGSEIEYGGAI VPSGEKLSPTKAIANAATVTIRQ 452
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 513 NVTIEQGRIVLRBKAKLSVNSLSQTCGS-LYHAGSTWDFVTPQPPQPPAANQILITLN 571
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 453 PAVLARGDLVLRDGVTVTFKDIQTQSPGRILMDGGTTLS-----AKEANLSLNG 501
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 572 LHLSSLSLANNNAVTPNPPAODSHPAVIGSTTAGSVTISGPIFFPFDLDDTAYDRYDM 631
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 LAVNLSSLDGTNKA-----ALKTEAADKNISLSGTTALIDTEGSPFENN- 546
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 632 LGSNOKINVLKQLQTKPPAN-----APSDITLGNEMPKYGYQGSWKLAMPDNTANNP 685
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 LKSASTYPLLEL---TTAGANTITLGALESTLTLQEPETHYQGNWQLSWANATSSK-- 601
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 686 YLKATWTYKYNCPERVASLVFNLSLWGLSILDIRSAHSATQASVDGRSYCRGLWVGVS 745
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 -IGSINWRTGTYP8PERKSNLPLNSLWGNFDIRSINQLIETKSSGSPFERELWSGIA 660
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 746 NFFYHDRDALGQGYRYISGGYSLGANSYP-GSSMFGLAPTEVFGSKDYVVCRSNHHACI 804
   :||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Db 241 -SFISNKAISFINN-----SVTATSAT-----CGAIYC-----SSTSAPKPV 277
QY 356 VSPDGEVVPFSSNVAAGKGGAIYAKLVANCGVPQVFLRNTAND-----GGAIVLGESG 410
Db 278 L7LSDNGELNFIQNTAITSGGAIYTDNLVSSGGPTLFXNNSAIDTAAPLGGAIADSG 337
QY 411 ELSLSADYGDIIIPDGNLKTAKENADVNGVTVSSQAIISMG--SGGKITTLRAKAGHOILF 469
Db 338 SLSLSALGDIDITPEGN---TVVKAS--SSQTTTNSINIGTNNAKIVQLRASQNTIYF 392
QY 470 NDIPIEMANGNNOPOASSKLLKINDGE-----GYTGDIVFA-----NGSSTLY 511
Db 393 YDPIIT-----TSITAALSDALNGLPDLAGNPAYQGTIVFSGEKLSEAAEAADNLKSTIQ 448
QY 512 QNVTIEQGRIVUREKAKLVNSLSQTGG--LYNEAGSTWDFVTPQPPQPPAANQLITLS 570
Db 449 QPLTLAGGQSLKSGVTLVAKSPSQSGPTLLMDAGTTLETADG-----ITIN 496
QY 571 NLHLSLSLLANNAVTPNPPAODSHPAVIGSTTAG--SVTISGPIPFEDLDDTDAYDRY 629
Db 497 NLVLNVDLSL-----KETKATLKATQASQTVLSSLSLVDPGNYEDV 541
QY 630 DMLGSNOKINVLKQLGTKPPANAP--SDLT---LGNEMPKYQYQGSWKLAWDPNTANNGP 685
Db 542 SW--NNPOV--FSCLTLTADDPANIHITDLAADPLEKNPIHMGYQGNWALSQEDTATKSK 598
QY 686 YTLKATWTKTGNPGERVASLVNLSWGSILDIRSAHSAIAQASVDGRSYCRGLWVGVS 745
Db 599 -AATLTWTKYGNPNPERRGTLVANTLWGSFVDVRSIQQLVATKVRSQOETRGICEGIS 657
QY 746 NPTHRDALCGQYRISIGSYISLGANSYFGS--SMFGLAPTEVGRSKDYVVCRSNHHACI 804
Db 658 NFFHKDSTKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKORDHPINKRASAYA 717
QY 805 GSVYLSQOALCG---SYLFG-----DAPIRASYGFGNQHKMTSYTFAEESDVRW 851
Db 718 ASHLQHLATLSPSLRYLPGSEBQVLFDAQI--SYIYSKNTMKTYTTPQAKGESSW 775
QY 852 DNCLAGEIGAGLPVITPISKLYLNLRLPFVQAEFSYADHESFTEBQDQ--ARAFKSGHLL 910
Db 776 YNDGCALELASSLPHTALSHEGLPHAYFPFIKVEASYIHQDSFKERNITLVRSPDSGLI 835
QY 911 NLSVPVGKFDRCSSGTHPNKYSFMAAYICDAYRTISGTETTLTSHOETWTTDAFLARHG 970
Db 836 NVSPVIGITFERFSRNERASYEATVIYVADVVRKPNPDCCTTALLINNTSWKTGTNLRSQA 895
QY 971 VVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
Db 896 GIGRAGIFYAFSPNLEVTLSMEIRGSSRSYNADLGKQFP 937

```

RESULT 30

```

US-09-198-452A-472
; Sequence 472, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 472
; LENGTH: 927
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...927
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-472

```

```

Query Match 18.9%; Score 994.5; DB 2; Length 927;
Best Local Similarity 29.0%; Pred. No. 1.1e-77;
Matches 308; Conservative 157; Mismatches 413; Indels 183; Gaps 38;
QY 1 MOTSFHKFPLSMILAYSCSLNGGYYAAEIMV---PQGIYDGE--TLTVSFPFVTVIGDPSG 56
Db 1 MKSSLHWFLLISSSLALP--LSLNFSAFAAVWEINLGTPTNSFSGFGTYT---PPAQTNDAG 56
QY 57 TTVPSAGELTLKUNLONSTAALPLSCFGNLGSFTVLGRGHSLTFFENIRSTSTNGAALSNSA 116
Db 57 TIYNTUTGDVSYNA--GSPALTASCCKETTGMLSPQGHGYQLLQNLIDAGAN--CIFTNTA 114
QY 117 ADGLFTIEBGFELSFNSCNLSLAVLPAATTNKGSOTPTTSTPSTNGTIYSKTDLLLNNE 176
Db 115 ANKLASFSGFSLSL-----IQTTNATTTGTAIKSTGACSIQSNY 154
QY 177 KFSFVSNLVSGDGGDAIDAKSLTVQGISKLCVQENTAOADGACQVTVTSFSAMANEAPIA 236
Db 155 SCYFQGNFSDNGGALQSSISL--SLNPNLTFAKNKATQKGGAL-----YST----- 200
QY 237 FVANVAGVRGGIAAQQDQCGVSSSTSTEDPVWFSRNTAVEFDGNVARVGGIYSYGN 296
Db 201 -----GGIT-----INNTLNS-----ASFSENTA-----ANNGAIYTEAS 231
QY 297 VAPLNNKCTL--FLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAIFCKNGAQAGSNNSGS 355
Db 232 -SFISNKAISFINN-----SVTATSAT-----GGAIVC-----SSTSAPKPV 268
QY 356 VSPDGEVVPFSSNVAAGKGGAIYAKLVANCGVPQVFLRNTAND-----GGAIVLGESG 410
Db 269 L7LSDNGELNFIQNTAITSGGAIYTDNLVSSGGPTLFXNNSAIDTAAPLGGAIADSG 328
QY 411 ELSLSADYGDIIIPDGNLKTAKENADVNGVTVSSQAIISMG--SGGKITTLRAKAGHOILF 469
Db 329 SLSLSALGDIDITPEGN---TVVKAS--SSQTTTNSINIGTNNAKIVQLRASQNTIYF 383
QY 470 NDIPIEMANGNNOPOASSKLLKINDGE-----GYTGDIVFA-----NGSSTLY 511
Db 384 YDPIIT-----TSITAALSDALNGLPDLAGNPAYQGTIVFSGEKLSEAAEAADNLKSTIQ 439
QY 512 QNVTIEQGRIVUREKAKLVNSLSQTGG--LYNEAGSTWDFVTPQPPQPPAANQLITLS 570
Db 440 QPLTLAGGQSLKSGVTLVAKSPSQSGPTLLMDAGT-----TLE 479
QY 571 NLHLSLSLLANNAVTPNPPAODSHPAVIGSTTAG--SVTISGPIPFEDLDDTDAYDRY 630
Db 480 TADGSLSIICSQCRFLK-----RQEXTLKATQASQTVLSSLSLVDPGNYEDVS 532
QY 631 WLGSNOKINVLKQLGTKPPANAP--SDLT---LGNEMPKYQYQGSWKLAWDPNTANNGPY 686
Db 533 W--NNPOV--FSCLTLTADDPANIHITDLAADPLEKNPIHMGYQGNWALSQEDTATKSK-- 588
QY 687 TLKATWTKTGNPGERVASLVNLSWGSILDIRSAHSAIAQASVDGRSYCRGLWVGVSVN 746
Db 589 AATLTWTKYGNPNPERRGTLVANTLWGSFVDVRSIQQLVATKVRSQOETRGICEGISN 648
QY 747 FFYHRRDALCGQYRISIGSYISLGANSYFGS--SMFGLAPTEVGRSKDYVVCRSNHHACIG 805
Db 649 FFKHSDSTKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKORDHPINKRASAYA 708
QY 806 SVYLSQOALCG---SYLFG-----DAPIRASYGFGNQHKMTSYTFAEESDVRWD 852
Db 709 SLHLQHLATLSPSLRYLPGSEBQVLFDAQI--SYIYSKNTMKTYTTPQAKGESSWY 766
QY 853 NNCLAGEIGAGLPVITPISKLYLNLRLPFVQAEFSYADHESFTEBQDQ--ARAFKSGHLLN 911
Db 767 NDGCALELASSLPHTALSHEGLPHAYFPFIKVEASYIHQDSFKERNITLVRSPDSGLIN 826
QY 912 LSPVVGKFDRCSSGTHPNKYSFMAAYICDAYRTISGTETTLTSHOETWTTDAFLARHG 971
Db 827 VSPVIGITFERFSRNERASYEATVIYVADVVRKPNPDCCTTALLINNTSWKTGTNLRSQA 886

```

Qy	972	VRGSMYASLTSNIEVYGHGRVEYRDASRGYGLSAGSRVRF	1012
		: : : : : : : : : : :	
Db	887	IGRAGIFYAFSPNLEVTNSLMEIRGSSRSYNADLGKQFQF	927

```

RESULT 31
US-09-438-185A-446
  Sequence 446, Application US/09438185A
  Patent No. 6822071
  GENERAL INFORMATION:
  ; APPLICANT: Stephens, Richard
  ; APPLICANT: Mitchell, Wayne
  ; APPLICANT: Kalman, Sue
  ; APPLICANT: Davis, Ronald
  ; APPLICANT: The Regents of the University of California
  TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
  FILE REFERENCE: 018941-000411US
  CURRENT APPLICATION NUMBER: US/09/438,185A
  PRIOR FILING DATE: 2002-03-13
  PRIOR APPLICATION NUMBER: US 60/108,279
  PRIOR FILING DATE: 1998-11-12
  PRIOR APPLICATION NUMBER: US 60/128,606
  PRIOR FILING DATE: 1999-04-08
  NUMBER OF SEQ ID NOS: 1074
  SOFTWARE: FastSeq for Windows Version 3.0
  SEQ ID NO 446
  LENGTH: 1414
  TYPE: PRT
  ORGANISM: Chlamydia pneumoniae
  FEATURE:
  ; OTHER INFORMATION: CFP0444
US-09-438-185A-446

```

Query Match 18.7%; Score 985.5; DB 2; Length 1414;
Best Local Similarity 28.0%; Pred. No. 1.4e-76;
Matches 302; Conservative 148; Mismatches 404; Indels 226;

Qy	51	IGDPSGTTVFSAGBLTKLNDLSI---AALPLSCFNGLLGSFTVLGRGHSITPFENIRTS	107
Db	443	VSDKTGLSIANNQEBVSLTSNAATVSGCAIYATKC-----TLTGNG-SLTFDGNAGT	493
Qy	108	NGRAALNSAADGLFTIEGFELSFSCNCSLLAVLPAATNKGSTPTTSTPSTNGTIYSK	167
Db	494	SGGAIYTTETD--FTLTG-----STGTVTFSTNKTAKTGALYSK	530
Qy	168	TDLLLLNNKFSFYSLNVSGD-----GGAI-----DAKSLTVQGI-----SKLCVF	208
Db	531	GNNSLSGNTWLLFSGNKATGPSNCSANQEGCGAILSLFSLASVSTKGLMIENENVSL	590
Qy	209	QENTAQDGGACQVVTFSAMANEAPAFVANVAGVGGGIAAVQDQGGVSSSSTSTEDP	268
Db	591	SGNTATVSGGA--IYATKCALHGNTLTTFDGNATARTAGAI-----YTETEDF	636
Qy	269	VVSFSRTAVFPDGNVARVGGIYSGNVAFLNNGKTLFLNNVASPVYIAKQPTSGQAS	328
Db	637	TLTGSTGT-VTFSTNTAKTALHTKGTSTPTKNKALVFGNSA-----TATAT	684
Qy	329	NTSNNGDYCGAIFC-----KNGAQAAGSNNSGSVS-----FDGEGV	363
Db	685	TTTTDQEGCGAILCNISESDIATKSLTLTENESLSPFNNTAKRSGGGIYAPKCVISGES	744
Qy	364	VPFSSNVAAGKGGAIYAKKLVSANCGVPQFLNTANDGGAIYLGESGELSISADYGDII	423
Db	745	INFNGNTATSGGAIYSKNLSITANGVPVFTNNSGGKGGAIYIADSGELSLEADTGDITP	804
Qy	424	DGNLKRITAKENAADVNGTVSSQAIMSGSGGKITTLRAKAGHOILFNDPIEM-----ANG--	478
Db	805	SGN---RATEGSTPN-----SIHLGAGAKITKLAAAPGHTIYFYDPTTWEAPASGGT	854
Qy	479	-----NNQPAQSSKLLKINDGEGVYTGDIV-----ANG	506
Db	955	IEELVINPVVKAIIVPPQPKNGPTIASVPVVPVAPANPTGTVTFSSGKLPSODASIPANT	914

Qy	507	SSTLYQNVITIEGRIVLREKAKLSVNSLSQTGS--LYMEAGSTWDFVTVPQPOPPAAHQ	565
Db	915	TTILNOKINLAGNVVLKEGATLQVYSPTQPDSTVFMGATITLTTTN-----NTDG	968
Qy	566	LITLSNHLHLSLSLANNNAVTPNPTPPAQDSHPAVIGSTTAGSVTSGPIFPEDLDDTA	625
Db	969	SIDLKNLVSNLVDALDGKRMIT-----IAVNSTSGLISGDLKFHNNEGSP	1014
Qy	626	YDRYDMLGSNQKINVLKQL--GTKPPAN-----APSDLTLGNEMPKYGVQGSWKLAWDP	678
Db	1015	YDN---PGLKANLNLPLDLSSPTGVNLDDFPNIPSSMA-----APDYGVQGSWTLV--P	1065
Qy	679	NTANNGPYTILKATWTKTYGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRG	738
Db	1066	KVGAGGKVTILVAEQALGYTPKPELRATLVPSNLWNYNIHSIQEIIATMSDADSHPG	1125
Qy	739	LWYSGVSNPFYHRRDALGOQYRYISGGYSLGANSYFGSS--MFLGAFTEVPGRSKDYVCCR	797
Db	1126	IWIGGIGNAFHQDKENAGFRLISRGYIVGGSMTTPQETTPAVAFSQLPFGSKDYVUSD	1185
Qy	798	SNHHACIGSYVLSTQALC--GSYLP-----GDATY-----RASLYGF	832
Db	1186	IKSQVYAGS-----LCAQSSYVITPLHSSLRHVLSKVLPPELPGETPLVLHGQVSYGR	1237
Qy	833	GNOHMTSYTFAESDVRWNNCLAGEIGAGLPVITPSKLYLNLRLPFVQAEPSYADHE	892
Db	1238	NHNMTTKLANNTQGSKDWDHSHFAVEVGGSVPVDL--NYRYLTSYSPYVKLQVSVNQK	1295
Qy	893	SFTEEGDQARAFKSGHLLNLSVPVGKFCRCSSTHPNKYSFMAAYICDAYRTISGTETTL	952
Db	1296	GFOEVAADPRIFDAHLNVVNSIPMGLTFKHESAKPFSALLTLGYAVDAYRDHPHCLITS-	1354
Qy	953	LSHOETWTTTDAFLARHGVVVVRGSMYASITSNIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
Db	1355	LWGTGWSVTFATNLSPQAPFAASGHLKLHLGLDCPAGSCGLRSSRSRYNANCGRYGF	1414

```

RESULT 32
US-09-438-185A-17
; Sequence 17, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 780
; TYPE: Prt
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPn0015
US-09-438-185A-17

```

Query Match 18.6%; Score 980.5; DB 2; Length 780;
Best Local Similarity 29.3%; Pred. No. 1.4e-76;
Matches 261; Conservative 142; Mismatches 320; Indels 167; Gaps 26;
Qy 175 NKFSPSYSLVSGDGAIDAKSLITVGISGLCVFQENTQAQDGGACQVVTFSFMANEAP 234
Db 6 NVSLILSKFNSTNGCAITAKTLISLTGTWTMSALFSENTSSKGGALQTSDLATITNGCE 65


```
Db 718 FRYASRPNVNVGILSRRLQNLPLMFHFLCAYGATDMKTDYANPMPVNSWRNCA 777
QY 858 GEIGAGLPVITPSPKLYLNLBPVQAFSPYADHESFTEBGQAPAPKSGHLLNLSPVG 917
Db 778 IECGSMPLLVFENGRLFGAIFPMKQLVYAYQGFKEETADGRFSGSLTSISVPLG 837
QY 918 VKPDRCSSTHPKYSFMAIYICDAVRTISGTETLLSHQETTTDAFLARHGVPVGRSM 977
Db 838 IRPEKALSDQVLYDFSFSYIPDIFRKPDSCEAALVISGDSMLVPAAHVSRHAFVGSOTG 897
QY 978 YASLTSNIEVYHGRYEYRDASRGYGLSAGSRVRF 1012
Db 898 RYHFNDEYELLRCGSIIECRPHARNYINCGSKFRF 932

RESULT 34
US-09-198-452A-15
; Sequence 15, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 15
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; NAME/KEY: SITE
; LOCATION: 1...922
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-15

Query Match 18.2%; Score 958; DB 2; Length 922;
Best Local Similarity 27.8%; Pred. No. 1.8e-74;
Matches 293; Conservative 145; Mismatches 381; Indels 236; Gaps 32;

QY 31 MYPQIYDGETLVSPFTVIGDPSGTTVFSAGELTLKNLNSIAALPLSCFNLGSGFT 90
Db 31 LTPEDSFHGDSONAERSYNV---QAGDVYSLTGDVSIQNSVDS--ALNKACFXVTSGSVT 85
QY 91 VLGRGHSLTPEIRITST--NGAALSNSAADGLFTIEGPKELSPSCNCSLLAVLPAATNK 148
Db 86 FAGNHGXYFNMISSGTTKEGAVLCCQDPQATARPSGFTLSF----- 128
QY 149 GSQPTPTTSTPSNGTIYKTDILLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVF 208
Db 129 -NQSP--GDIEQGLYKSNALMLNNYVVRPEQNSKTKGALSGANVTIVGNVDSVSP 185
QY 209 QENTAAQAGCAGQVVTSPSAMANEAPIAFVANAGVGGGIAAVQDQGGVSSSTSTEDP 268
Db 186 YQNAATFGG-----AIHSGPLQIAVQAIR----- 212
QY 269 VVSFRNTAVEPDGNVARVGGIYSGNVAFNLNGKTLPLNNVASPVVIAAKQPTSGQAS 328
Db 213 ---FAQNTAK--NGS-----GGALYSDGDDIDQDAYLVFREALTAT-- 252
QY 329 NTSNNYDGGGAIFCKNGAQAGSNNSGS-----VSFDGEGVVFSSNVAAGKGAIYAKK 382
Db 253 -----GKGAVCC-----LPTSGSSTPPVIVTFSDNKQLVPERNHSIMCGGAIYAKK 299
QY 383 LSVANGCPVQFLRNIA-----NDGAIYLGSGGELSLGADYGDIIIPDGNLKRKTAKENAD 437
Db 300 LSISSGGPTLFINNISYANSQNLGAIAIDTGGGEISLSAEGKTTITFOGN-----RTSLPF 354
QY 438 VNGVTVSQOAIMSGSGKITTILRAKAGHOILFNDPIEHANGNNQPAQSSKLLKLN---DG 494
```

```
355 LNGIHLLQNA-----KFLKLQARNGSYSTFYDPI-----TSEADSGSTQININGDPKN 401
QY 495 EGYTGDIIVA-----NGSSTLYQNVTIEQGRIVLRREKAKLSVNSLSQTGGS--LYM 543
Db 402 KEYTGTILFSCSEKSLANDPRDFKSTIPQNVNLSAGYLVIKEGAEVTVSKPTQSPGSHLV 461
QY 544 BAGSTWDFVTQPQQPPAANQLITLSNLHLSLSSLLANNAVTPPTNPPPAQDSHPAIVG 603
Db 462 DLGT-----KLITASKEDIAITGLAIDIDSL-----SSSSTAATAVIK 496
QY 604 STTAG--SVTISGPIP-----PEDLDDTAYDRYDMLGSNQKINVLKLGITKPPANAP 654
Db 497 ANTANKQLSVTDSIELISPTGNAYEDLR-----MRNSQTTPPLLSLE-----PGAGS 543
QY 655 SDLTLGNEM---PKYGYQGSNKLAWDPNTANNNGPYTLKATWTKTGYNPGPBRVASLVPNS 711
Db 544 VVTAGDPLPVSPHYGPGQGNKLAWTGKNKVGEP---PMDKINYKPRPEKGNLVN 599
QY 712 LWSGILDIRSAHSAIQASVDGRSYCRGLWVSGVGNFPYHDDRDLALGOQGYRISGYSGLAN 771
Db 600 LMGNAVDVRSIMQVQETHASSLQTDRLGWLIDGIGNLFHVSASEDNIRYHNSGGYVLSVN 659
QY 772 SYFGSSMF-GIAFTEVFGRSKDVVVCRSNHHACIGSVVLSTQQALCGSYLF-----GDA 824
Db 660 NEITPKHTSMFSQLPSRDKDYAVSNNEYR-----MYL-----GSYLYQYTTSLGNI 707
QY 825 FIRAS-----YFGNQHMKTSTYTFABESDVRDNNCLA 857
Db 708 FRYASRPNVNVGILSRRLQNLPLMFHFLCAYGATDMKTDYANPMPVNSWRNCA 767
QY 858 GEIGAGLPVITPSPKLYLNLBPVQAFSPYADHESFTEBGQAPAPKSGHLLNLSPVG 917
Db 768 IECGSMPLLVFENGRLFGAIFPMKQLVYAYQGFKEETADGRFSGSLTSISVPLG 827
QY 918 VKPDRCSSTHPKYSFMAIYICDAVRTISGTETLLSHQETTTDAFLARHGVPVGRSM 977
Db 828 IRPEKALSDQVLYDFSFSYIPDIFRKPDSCEAALVISGDSMLVPAAHVSRHAFVGSOTG 887
QY 978 YASLTSNIEVYHGRYEYRDASRGYGLSAGSRVRF 1012
Db 888 RYHFNDEYELLRCGSIIECRPHARNYINCGSKFRF 922

RESULT 35
US-09-438-185A-15
; Sequence 15, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0013
US-09-438-185A-15

Query Match 17.8%; Score 939.5; DB 2; Length 866;
Best Local Similarity 27.0%; Pred. No. 6.7e-73;
```


QY 479 -----NNPAQSSKLLKINDGEGYTGDIIVFA-----ANGSTLYQNVITREQ 519
 Db 730 VPPPOKNGPIASVPVAPANPNTGTVFSSGKLPSQDASIPANTTILNOKINLAGG 789
 QY 520 RIVLRKAKLSVNSLSQTS-LYMEAGSTWDFVTPPOPPQPPAANQLITLSNHLSS 578
 Db 790 NVVLAKGATLQVYSTQPPDSVFMADGTTLETTTN-----NTDGSIDLKNSLVNDA 843
 QY 579 LLANNAVTPNPPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDATAVDRYDMLGNSOKI 638
 Db 844 LDGKEMIT-----IAVNSTGGGLKISGDLKFHNNEGSFYDN---PGLKANL 886
 QY 639 NVLKLQ-LGTRKPPAN-----APSDTLGNEMPKYGYGSKWKLAWDPNTANNPPTLKAT 691
 Db 887 NLPLFDLSTSGTVNLDNFPNPSMA-----APDYGYGSGMTLV--PKVAGGKVTLVAE 940
 QY 692 WTKTGNPGPERSVASLVPNSLWMSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHD 751
 Db 941 WQALGYTPKPELRATLVNLSLWNAVNIHSIQEITATMSDAPSHPGIWIIGIGNAFQD 1000
 QY 752 RDALGQGYRIISGGYSLGANSYFGSS-MFGLAFTVFGRSKDYVVCRSNHHACIGSVYLS 810
 Db 1001 KOKENAGFLLISRGYIVGSGMTTPQEXTPAVAFSOLFGKSKDYVVSIDIKSOVYAGS---- 1056
 QY 811 TQALC--GSYLP-----GDAFI-----RASVGFNGHMKTSYTFAE 845
 Db 1057 ----LCAQSSYVILPHSSLRHRLSVKLPPLPGTPLVHQVSYGRNHNMTTKLANNT 1112
 QY 846 ESDVRWDNN 854
 Db 1113 QKQSDWDH 1121

RESULT 37

US-09-438-185A-451
 ; Sequence 451, Application US/09438185A
 ; Patent No. 6822071
 ; GENERAL INFORMATION:
 ; APPLICANT: Stephens, Richard
 ; APPLICANT: Mitchell, Wayne
 ; APPLICANT: Kalman, Sue
 ; APPLICANT: Davis, Ronald
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
 ; FILE REFERENCE: 018941-00041105
 ; CURRENT APPLICATION NUMBER: US/09/438,185A
 ; PRIORITY FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: US 60/108,279
 ; PRIOR FILING DATE: 1998-11-12
 ; PRIOR APPLICATION NUMBER: US 60/128,606
 ; PRIOR FILING DATE: 1999-04-08
 ; NUMBER OF SEQ ID NOS: 1074
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 451
 ; LENGTH: 634
 ; TYPE: PRF
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; OTHER INFORMATION: Cpn0449
 US-09-438-185A-451

Query Match 15.3%; Score 808; DB 2; Length 634;
 Best Local Similarity 30.8%; Pred. No. 1.3e-61;
 Matches 215; Conservative 107; Mismatches 265; Indels 110; Gaps 19;

QY 358 FDGEVVPFSSNVAAGKGAIVAKKLSVANCPCVPQFLRNANDGGAIVLGSGELSLSD 417
 Db 6 FRGGGGSFNNIVQG-----TTAGNGGAISILAAEGCSLSAE 43
 QY 418 YGDIIFDGN-----LKRITAKENADVNGVTSSQAISMGSGGKITTLRAKAGHQLFNDDP 472
 Db 44 AGDITFNGNAIVATTPQTTKNSIDI-----GSTAKITNLRAISGHSIFFYDIP 91

QY 473 IEMANGNNQPAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNV 514
 Db 92 ITA-----NTAADSTDTLNLAKADAGNSTDYSGSVFSGEKLSEDEAKVADNLTLTKOPV 147
 QY 515 TIEQGRIVLRKAKLSVNSLSQT--GGSLYMEAGSTWDFVTPPOPPQPPAANQLITLSNHL 573
 Db 148 TLTAGNLVLKRGVTLDTKGTQTAGSSVIMDAGTTL-----KASTEETVLTGLS 196
 QY 574 LSLSSLLANNAVTPNPPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDATAVDRYDMLG 633
 Db 197 IPVDSL-----GEGKKVVIAASAASKNVALSGPILLDDNOGNAYENHD--LG 241
 QY 634 SNQKINVLKQ-LGTRKPPANAPSDTLGNEMPKYGYGSKWKLAWDPNTANNPPTLKAT- 691
 Db 242 KTODFSFVQLSALGATATTTDPAVPTVATP--THYGYGQGTWGMTVDDTAST--PKTKTATL 299
 QY 692 -WTKTGNPGPERSVASLVPNSLWMSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH 750
 Db 300 ANWTGVLNPEROCPLVPNSLWMSFSDIQAIQGVIERALSATLCSDRGFMAAGVANFLDK 359
 QY 751 DRDALGQGYRIISGGYSL-GANSYFGSSMFGLAFTVFGRSKDYVVCRSNHHACIGSVYL 809
 Db 360 DKKEKRYRHKSGYAIAGAAQTCSENLSIFAFQCLFGSKDPLVAKNHTDTYAGAFYI 419
 QY 810 S-----TQALCGSYLFGDAFI--RASVGFNGHMKTSYTFAEESDVRWDNNCL 856
 Db 420 QHITCSGFGICGLADKLPGSMHKLPLVLEQLAYSHVSNDLTKYTAYPEYKSGMGNAP 479
 QY 857 AGEIGAGLPITVTPSKLY-LNELRPPFVQABFSPYADHESPTTEGQARAPKSGHLLNLSVP 915
 Db 480 NMLGASHSY--PEYLHCFDTPYPIKLNUTYIRQDSFSEKTEGSGFDDSNLFLSLP 537
 QY 916 VGKFDRCSTHPNKYSFMAAYICDAYRTISGTTTLLSHQETWTTDAFHLARHGUVVRG 975
 Db 538 IGKFKFSDCNDPSYDLTSLSYVPLRNDPKCTTALVISGASWETVANNLARQALQVRA 597
 QY 976 SMYASLTSNIEVYCHGYREYRDASRGYLSAGSVRP 1012
 Db 598 GSHYAPSPMEVLGQPFVFEVRGSSRIYNVDLGGKQF 634

RESULT 38

US-09-198-452A-474
 ; Sequence 474, Application US/09198452A
 ; Patent No. 6559294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 474
 ; LENGTH: 643
 ; TYPE: PRF
 ; ORGANISM: Chlamydia pneumoniae
 US-09-198-452A-474

Query Match 15.3%; Score 804; DB 2; Length 643;
 Best Local Similarity 30.8%; Pred. No. 3.1e-61;
 Matches 214; Conservative 107; Mismatches 264; Indels 110; Gaps 19;

QY 360 GEGVVPFSSNVAAGKGAIVAKKLSVANCPCVPQFLRNANDGGAIVLGSGELSLSDY 419
 Db 17 GGGGGSFNNIVQG-----TTAGNGGAISILAAEGCSLSAEG 54
 QY 420 DIIFDGN-----LKRITAKENADVNGVTSSQAISMGSGGKITTLRAKAGHQLFNDDP 474
 Db 55 DITFNGNAIVATTPQTTKNSIDI-----GSTAKITNLRAISGHSIFFYDIP 102
 QY 475 MANGNNQPAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNV 516

Db	103	A-----NTAADSDTLNLKADAGNSTDYSGSVFSEKLSSEDAKVADNLTLAQPVTL	158
Qy	517	EGRIVLRKAKLSVNSLSQT--GGSLYMEAGSTWDFVTPQPPQPPAANQLTLNLHLUS	575
Db	159	TAGNLVLRKGVTLDTKGTQTAGSSVINDAGTTL-----KASTEETVLTGLSIP	207
Qy	576	LSLLANNVTPPTNPQADSHPAVIGSTTAGSVTISGPIPFEDLDDTAYDRYDLGNS	635
Db	208	VDLS-----GEGKKVJAASAARKNVALSGPILLDNQGNAYENHD--LGKT	252
Qy	636	QKINVLKQ--LGTKPPANAPSOLTLGNEMPKYQGSWKLAWDPNTANNPPTLTKAT--W	692
Db	253	QDPSFVOLSALGTATTDVPAVPTATP--THGYQGTWGTWVDDTAST--PKTKATLAW	310
Qy	693	TKTGYNPGERVASLPVNSLWMSILDIRSAHSAIQASVDGRSCYRCGLWVGVSNNPFYHDR	752
Db	311	TNTGYLPNERQGPLVPNSLWMSFSDIQAIQGVIERALSALTCLSDRGFAAGVAFDKDK	370
Qy	753	DALGQGYRVISGYSL--GANSYFGSSMFGLAFTVFRGSKDYVVCRSNHHACIGSVYLS-	810
Db	371	KGEKRYRKHSGGYATGGAAQTCSENLSIFAPCOLFGSKDPLVAKNHTDTTAGAFYIOH	430
Qy	811	-----TQALCGSYLFGDAFI--RASVGFQGNHMKTSYTPAEESDVRMDNNCLAG	858
Db	431	ITECSPFIGCLLDKLPQSGSHKPLVLEGQLAYSHVNDLTKTYTAYPEVKGSGWNAFNM	490
Qy	859	EIGAGLPPIVITPSKLY--LNELRPFVQAFPSYADHESFTREGDQARAFKSHLLNLNVPG	917
Db	491	MLGASSHSY--PEYLHCFTDYAPVYKLNLTYYIQDSSFSEKGTGRGFSDDSNLNLSPIG	548
Qy	918	VKDRCSSTHPNKYSFMAAYICDAYRTISGTETILLSHQETWTTDAFHLARHGVVVRGSM	977
Db	549	VKPEKFSDCNDFSJDLTLVSVPDLIRNDPKCTTALVISGASWETYNANLARQALQVRAGS	608
Qy	978	YASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF	1012
Db	609	HYAFSPMFEVLQGFVFEVRGSSRIYNDLGGKQF	643

RESULT 39
US-09-556-877-175
; Sequence 175, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 175
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(880)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-556-877-175

Query Match 14.6%; Score 769.5; DB 2; Length 880;
Best Local Similarity 26.2%; Pred. No. 5.5e-58;
Matches 264; Conservative 146; Mismatches 395; Indels 203; Gaps 33;
Qy 53 DPSGTVFSAGELTKNLNDNSIALPLSCFQNL-----GSPFVLGRGHSITFE 101

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:31:04 ; Search time 168 Seconds
(without alignments)
2516.924 Million cell updates/sec

Title: US-10-701-844-2
Perfect score: 5267
Sequence: 1 MQTSFHKFLSMILAYSCS.....VEYRDASRGYGLSAGSRVRF 1012

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5267	100.0	1012	4 US-10-701-844-2	Sequence 2, Appli
2	5267	100.0	1012	4 US-10-766-711-2	Sequence 2, Appli
3	5267	100.0	1012	5 US-10-931-779-2	Sequence 2, Appli
4	5131.5	97.4	1013	4 US-10-467-534-9	Sequence 9, Appli
5	5131.5	97.4	1013	5 US-10-498-327-5	Sequence 5, Appli
6	5131.5	97.4	1013	5 US-10-498-327-81	Sequence 81, Appli
7	5125.5	97.3	1013	5 US-10-498-327-83	Sequence 83, Appli
8	5123.5	97.3	1013	5 US-10-498-327-9	Sequence 9, Appli
9	5123.5	97.3	1013	5 US-10-498-327-13	Sequence 13, Appli
10	5123.5	97.3	1013	5 US-10-498-327-93	Sequence 93, Appli
11	5121.5	97.2	1013	5 US-10-498-327-1	Sequence 1, Appli
12	5121.5	97.2	1013	5 US-10-498-327-3	Sequence 3, Appli
13	5116	97.1	984	4 US-10-701-844-43	Sequence 43, Appli
14	5116	97.1	984	4 US-10-766-711-43	Sequence 43, Appli
15	5116	97.1	984	5 US-10-931-779-43	Sequence 43, Appli
16	5112.5	97.1	1013	5 US-10-498-327-17	Sequence 17, Appli
17	5106.5	97.0	1013	4 US-10-701-844-15	Sequence 15, Appli
18	5106.5	97.0	1013	4 US-10-701-844-16	Sequence 16, Appli
19	5106.5	97.0	1013	4 US-10-766-711-15	Sequence 15, Appli
20	5106.5	97.0	1013	4 US-10-766-711-16	Sequence 16, Appli
21	5106.5	97.0	1013	5 US-10-931-779-15	Sequence 15, Appli
22	5106.5	97.0	1013	5 US-10-931-779-16	Sequence 16, Appli
23	5090	96.6	1006	3 US-08-841-132-190	Sequence 190, App
24	5090	96.6	1006	5 US-10-872-155-190	Sequence 190, App
25	5084	96.5	982	3 US-09-841-132-176	Sequence 176, App
26	5084	96.5	982	5 US-10-872-155-176	Sequence 176, App
27	3336.5	63.3	670	5 US-10-197-220-169	Sequence 169, App

28	3336.5	63.3	670	6 US-11-109-468-169	Sequence 169, Appl
29	2552	48.5	505	4 US-10-701-844-17	Sequence 17, Appl
30	2552	48.5	505	4 US-10-766-711-17	Sequence 17, Appl
31	2552	48.5	505	5 US-10-931-779-17	Sequence 17, Appl
32	2350	44.6	458	4 US-10-701-844-36	Sequence 36, Appl
33	2350	44.6	458	4 US-10-766-711-36	Sequence 36, Appl
34	2350	44.6	458	5 US-10-931-779-36	Sequence 36, Appl
35	1735	32.9	325	4 US-10-701-844-37	Sequence 37, Appl
36	1735	32.9	325	4 US-10-766-711-37	Sequence 37, Appl
37	1735	32.9	325	5 US-10-931-779-37	Sequence 37, Appl
38	1365.5	25.9	973	4 US-10-312-273-3	Sequence 3, Appli
39	1359.5	25.8	973	4 US-10-352-618-2	Sequence 2, Appli
40	1128	21.4	928	3 US-09-428-122-2	Sequence 2, Appli
41	1122	21.3	926	3 US-09-738-269-57	Sequence 57, Appl
42	1122	21.3	926	4 US-10-023-437-57	Sequence 57, Appl
43	1091	20.7	928	4 US-10-312-273-95	Sequence 95, Appl
44	1090	20.7	928	4 US-10-312-273-115	Sequence 115, App
45	1090	20.7	928	4 US-10-282-122A-54681	Sequence 54681, A

ALIGNMENTS

RESULT 1
US-10-701-844-2
; Sequence 2, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-2

Query Match	100.0%;	Score	5267;	DB	4;	Length	1012;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1012;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MQTSFHKFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF	60				
Db	1	MQTSFHKFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF	60				
QY	61	SAGELTKMLDINSIAALPLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSNSAADGL	120				
Db	61	SAGELTKMLDINSIAALPLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALSNSAADGL	120				
QY	121	FTIEGPKELSFNSCNLSLLAVLPAATNKGSTPTTTSTPSNGIYISKTDLLLLNNEKFSF	180				
Db	121	FTIEGPKELSFNSCNLSLLAVLPAATNKGSTPTTTSTPSNGIYISKTDLLLLNNEKFSF	180				
QY	181	YSLNLSGDDGAIIDAKSLTVQGISKLCVQENTAAQDGGACQVVTFSFMANEAPIAFVAN	240				
Db	181	YSLNLSGDDGAIIDAKSLTVQGISKLCVQENTAAQDGGACQVVTFSFMANEAPIAFVAN	240				
QY	241	VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSFRNTAVEPDGNVARVGGGIYSYGNVAF	300				
Db	241	VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSFRNTAVEPDGNVARVGGGIYSYGNVAF	300				
QY	301	NNGKTLFLNNVASPVYIAAKQPTSGQASNTNNYDGGGAIFCXKNGAQAQAGNSGVSFDG	360				
Db	301	NNGKTLFLNNVASPVYIAAKQPTSGQASNTNNYDGGGAIFCXKNGAQAQAGNSGVSFDG	360				


```
QY 361 EGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSSELSSADYGD 420
Db 361 EGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSSELSSADYGD 420
QY 421 IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN 480
Db 421 IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN 480
QY 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS 540
Db 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS 540
QY 541 LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA 600
Db 541 LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA 600
QY 601 VIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNOKINVLKQLGTKPPANAPSDLTG 660
Db 601 VIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNOKINVLKQLGTKPPANAPSDLTG 660
QY 661 NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLMSILDIR 720
Db 661 NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLMSILDIR 720
QY 721 SAHSAIQASVDGRSYCRGLMWVSGVSNFFYHDDRDLGQGYRIVISGGYSLGANSYFGSSMFG 780
Db 721 SAHSAIQASVDGRSYCRGLMWVSGVSNFFYHDDRDLGQGYRIVISGGYSLGANSYFGSSMFG 780
QY 781 LAPTEVFGRSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA 840
Db 781 LAPTEVFGRSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA 840
QY 841 YTPAESDVRWNNCLAGEIGAGLPIVITPSKLYLNLRLPFFVQAEFSYADHESFTEEGDQ 900
Db 841 YTPAESDVRWNNCLAGEIGAGLPIVITPSKLYLNLRLPFFVQAEFSYADHESFTEEGDQ 900
QY 901 ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960
Db 901 ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960
QY 961 TDAFHARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012
Db 961 TDAFHARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012

RESULT 2
US-10-766-711-2
; Sequence 2, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; PRIOR FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-2
Query Match 100.0%; Score 5267; DB 4; Length 1012;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOTSPHKFFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTVTF 60
```

```
Db 1 MOTSPHKFFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTVTF 60
QY 61 SAGELTLKLNLDNSAALPLSCFGLNLGSPFTVLGRGHSLTFTENIRTSNGAALSNSAADGL 120
Db 61 SAGELTLKLNLDNSAALPLSCFGLNLGSPFTVLGRGHSLTFTENIRTSNGAALSNSAADGL 120
QY 121 FTIEGPKELSPSNCNSLLAVLPAATTKGSGTPTTSTPNSGTIYSKTDLLLLNNKPSF 180
Db 121 FTIEGPKELSPSNCNSLLAVLPAATTKGSGTPTTSTPNSGTIYSKTDLLLLNNKPSF 180
QY 181 YSNLVSGDGGDAIDAKSLTVQGISKLCVPOENTAQADGACQVVTSPSAMANEAPIAVAN 240
Db 181 YSNLVSGDGGDAIDAKSLTVQGISKLCVPOENTAQADGACQVVTSPSAMANEAPIAVAN 240
QY 241 VAGVGGGIIAAVDQGGQGVSSSTSTEDPVVSFSRNTAVEFDGNVARGVGGIYSYGNVAF 300
Db 241 VAGVGGGIIAAVDQGGQGVSSSTSTEDPVVSFSRNTAVEFDGNVARGVGGIYSYGNVAF 300
QY 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGGAIFCXGGAQAGSNNSGSVFDG 360
Db 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGGAIFCXGGAQAGSNNSGSVFDG 360
QY 361 EGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSSELSSADYGD 420
Db 361 EGVVFPSSNVAAGKGGAIYAKKLSVANGCPVQFLNIANDGGAIVLGSSELSSADYGD 420
QY 421 IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN 480
Db 421 IIFDGNLKRKTAKENAADVNGVTSSQALSMGSGGKIITTLRAKAGHQLIIFNDPIEMANGNN 480
QY 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS 540
Db 481 QPAQSSKLLKINDGEGYTDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGGS 540
QY 541 LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA 600
Db 541 LYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPQAQDSHPA 600
QY 601 VIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNOKINVLKQLGTKPPANAPSDLTG 660
Db 601 VIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNOKINVLKQLGTKPPANAPSDLTG 660
QY 661 NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLMSILDIR 720
Db 661 NEMPYGYQGSWKLAWDNTANNNGPYTLKATWTGTGYNPGERVASLVPNSLMSILDIR 720
QY 721 SAHSAIQASVDGRSYCRGLMWVSGVSNFFYHDDRDLGQGYRIVISGGYSLGANSYFGSSMFG 780
Db 721 SAHSAIQASVDGRSYCRGLMWVSGVSNFFYHDDRDLGQGYRIVISGGYSLGANSYFGSSMFG 780
QY 781 LAPTEVFGRSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA 840
Db 781 LAPTEVFGRSKDYVVCRSNHHACIGSVYLSLSSLLANNAVTPNPPQAQDSHPA 840
QY 841 YTPAESDVRWNNCLAGEIGAGLPIVITPSKLYLNLRLPFFVQAEFSYADHESFTEEGDQ 900
Db 841 YTPAESDVRWNNCLAGEIGAGLPIVITPSKLYLNLRLPFFVQAEFSYADHESFTEEGDQ 900
QY 901 ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960
Db 901 ARAFKSGHLLNLVSPGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960
QY 961 TDAFHARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012
Db 961 TDAFHARHGVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012

RESULT 3
US-10-931-779-2
; Sequence 2, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
```

APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-2

Query Match 100.0%; Score 5267; DB 5; Length 1012;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQTSPHKPFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60

QY 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNNGAALSNSAADGL 120
DB 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNNGAALSNSAADGL 120

QY 121 FTIEGPKELSFNCNSLLAVLPAATTNKGSTPTTSTPSNGTIYKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSFNCNSLLAVLPAATTNKGSTPTTSTPSNGTIYKTDLLLNNEKPSF 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAOAGGACQVTVSFSAMANEAPIAFVAN 240
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAOAGGACQVTVSFSAMANEAPIAFVAN 240

QY 241 VAGVRGGIAAVODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVGGGIYSGNVAF 300
DB 241 VAGVRGGIAAVODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVGGGIYSGNVAF 300

QY 301 NNGKTLFLNNVSPVIAAKQPTSGQASNTSNNGYDGGAIFCXNGAAGSNNSGSVSFDG 360
DB 301 NNGKTLFLNNVSPVIAAKQPTSGQASNTSNNGYDGGAIFCXNGAAGSNNSGSVSFDG 360

QY 361 EGVVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 420
DB 361 EGVVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 420

QY 421 IIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLLRAKAGHQLFPNDPIEMANGNN 480
DB 421 IIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLLRAKAGHQLFPNDPIEMANGNN 480

QY 481 QPAQSSKLLKINDGEGYTGDI VFAANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGS 540
DB 481 QPAQSSKLLKINDGEGYTGDI VFAANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGS 540

QY 541 LYMEAGSTWDFVTPQPPQPPAANQLITLNLHLSSLANNVTPPNTPPAODSHPA 600
DB 541 LYMEAGSTWDFVTPQPPQPPAANQLITLNLHLSSLANNVTPPNTPPAODSHPA 600

QY 601 VIGSTTAGSVTISGPIPEDLDTPAYDRYDWLGSNQKINVLKQLGKTPPANAPSDLTG 660
DB 601 VIGSTTAGSVTISGPIPEDLDTPAYDRYDWLGSNQKINVLKQLGKTPPANAPSDLTG 660

QY 661 NEMP KYQGSWKLAWDPNTANNPYTLKATWTGTGNPGERVASLVPNSLMGSIILDIR 720
DB 661 NEMP KYQGSWKLAWDPNTANNPYTLKATWTGTGNPGERVASLVPNSLMGSIILDIR 720

QY 721 SAHSAIQASVDGRS YCRGLWVSGVSNPFYTHDRDALGQGYRIYSGYSLGANSYFGSSMFG 780
DB 721 SAHSAIQASVDGRS YCRGLWVSGVSNPFYTHDRDALGQGYRIYSGYSLGANSYFGSSMFG 780

QY 781 LAPTEVFGRSKYDYVCRSNHHACIGSVYLSQOALCGSYLFGDAFIRASYGFGNOHMKTS 840

DB 781 LAPTEVFGRSKYDYVCRSNHHACIGSVYLSQOALCGSYLFGDAFIRASYGFGNOHMKTS 840

QY 841 YTPAESDVRWNNCLAGEIGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTEEGDQ 900

DB 841 YTPAESDVRWNNCLAGEIGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTEEGDQ 900

QY 901 ARAFKSGHLLNLSEVPVGVKFDRCSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960

DB 901 ARAFKSGHLLNLSEVPVGVKFDRCSTHPNKYSFMAAYICDAYRTISGTETLLSHQETWT 960

QY 961 TDAPFLARHGUVVRGSMYASLTSNIETVYGHGREYRDASRGYGLSAGSRVRF 1012

DB 961 TDAPFLARHGUVVRGSMYASLTSNIETVYGHGREYRDASRGYGLSAGSRVRF 1012

RESULT 4
US-10-467-534-9
; Sequence 9, Application US/10467534
; Publication No. US20040131625A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verilant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/10/467,534
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-467-534-9

Query Match 97.4%; Score 5131.5; DB 4; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSPHKPFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCSLNGGGYAAEIMVPOGGIYDGETLTVSPYTVIGDPSGTTVF 60

QY 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNNGAALSNSAADGL 120
DB 61 SAGELTLKNDLSIAALPLSCFNGMLGSGFTVLGRGHSLTFENIRTSNNGAALSNSAADGL 120

QY 121 FTIEGPKELSFNCNSLLAVLPAATTNKGSTPTTSTPSNGTIYKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSFNCNSLLAVLPAATTNKGSTPTTSTPSNGTIYKTDLLLNNEKPSF 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAOAGGACQVTVSFSAMANEAPIAFVAN 240
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAOAGGACQVTVSFSAMANEAPIAFVAN 240

QY 241 VAGVRGGIAAVODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVGGGIYSGNVAF 300
DB 241 VAGVRGGIAAVODGQGGVSSSTSTEDPVVPSFRNTAVFDGNVAVGGGIYSGNVAF 300

QY 301 NNGKTLFLNNVSPVIAAKQPTSGQASNTSNNGYDGGAIFCXNGAAGSNNSGSVSFD 359
DB 301 NNGKTLFLNNVSPVIAAKQPTSGQASNTSNNGYDGGAIFCXNGAAGSNNSGSVSFD 359

QY 360 GEGVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 419
DB 360 GEGVFPSSNVAAGKGAIAKLSVANCQGVQFLRNANDGGAIYLGESGELSADYD 419

QY 420 DIIFDGNLKRKTAKENADVNGTVSSQAI SMGSGGKITTLLRAKAGHQLFPNDPIEMANGN 479

421	Db	DIIFDGNLXRTAKENAADVNGVTVSSQASISWMSGGKITTFRAKAGHOILFNDPIEMANGN	480
480	Qy	NQPAQSSKLLKJNDGSGYTGDIIVFANGSSTLYQNVTTIEQRIVLREKAKLSVNSLSQTCG	539
481	Db	NQPAQSSPELJNDGSGYTGDIIVFANGSSTLYQNVTTIEQRIVLREKAKLSVNSLSQTCG	540
540	Qy	SLYWEAGSTWDFVTPOPPQOPPAANQLITLSNLHLSLSSLLANNAVTPPTNPPAQDSHP	599
541	Db	SLYWEAGSTLDFVTPOPPQOPPAANQLITLSNLHLSLSSLLANNAVTPPTNPPAQDSHP	600
600	Qy	AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNQKINVLKLOLQCTKPPANAPSDLT	659
601	Db	AIIIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSNQKIDVLKLOLQCTQSANAPSDLT	660
660	Qy	GNEMPKYGYQGSWKLAWDPNTANNGPYTLKATWTNTGYNPGPERSVSLVPNSLWGSILDI	719
661	Db	GNEMPKYGYQGSWKLAWDPNTANNGPYTLKATWTNTGYNPGPERSVSLVPNSLWGSILDI	720
720	Qy	RSAHSAIQASVDRGSRYCRGLWVSGVSNFYHDDRDLGQGYRTYISGGYSLGANSYFGSSMF	779
721	Db	RSAHSAIQASVDRGSRYCRGLWVSGVSNFYHDDRDLGQGYRTYISGGYSLGANSYFGSSMF	780
780	Qy	GLAFTVEFGKSDYVVCVRNNHACIGSVYLSQOALCGSYLFGDAPFIRASVYGFNGOHMKT	839
781	Db	GLAFTVEFGKSDYVVCVRNNHACIGSVYLSQOALCGSYLFGDAPFIRASVYGFNGOHMKT	840
840	Qy	SYTFAEESDVRDNNCLAGEIGAGLPIVITPISKLYLNLRLPPFVQAEFSYADHESFTEEGD	899
841	Db	SYTFAEESDVRDNNCLVGEIGVGLPIVITPISKLYLNLRLPPFVQAEFSYADHESFTEEGD	900
900	Qy	QARAFKSGHLNLNLSPVGVKPFDRCSSTHPNKYSFMAAYICDAYTISGTETTLTLLSHQETW	959
901	Db	QARAFKSGHLNLNLSPVGVKPFDRCSSTHPNKYSFMAAYICDAYTISGTETTLTLLSHQETW	960
960	Qy	TTDAPFLAHGUVVRGSMYASLTSNIEYVGHGRYEYDRDAQSGYLSAGSVRPF	1012
961	Db	TTDAPFLAHGUVVRGSMYASLTSNIEYVGHGRYEYDRDAQSGYLSAGSVRPF	1013

```

RESULT 5
US-10-498-327-5
  Sequence 5, Application US/10498327
  Publication No. US20050106162A1
  GENERAL INFORMATION:
  APPLICANT: Grandi, Guido
  APPLICANT: Ratti, Giulio
  TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
  FILE REFERENCE: 002441.00085
  CURRENT APPLICATION NUMBER: US/10/498,327
  CURRENT FILING DATE: 2004-06-10
  PRIOR APPLICATION NUMBER: PC7/IB02/05761
  PRIOR FILING DATE: 2002-12-12
  NUMBER OF SEQ ID NOS: 262
  SOFTWARE: PatentIn version 3.2
  SEQ ID NO 5
  LENGTH: 1013
  TYPE: PRT
  ORGANISM: Chlamydia trachomatis
US-10-498-327-5

```

Qy	121	FTIEGPKELSF	SNCSNLLAVLPAAT	NNKGSQTPTT	STPSNGTIYSKTDLL	LLNNEKFSF	180		
Db	121	FTIEGPKELSF	SNCSNLLAVLPAAT	NNNGSQTPTT	STPSNGTIYSKTDLL	LLNNEKFSF	180		
Qy	181	YSNLVSGDGGAI	DAKSLTVQGISKL	CVPOBNTAODGCA	QCVVTSFSA	MANEAPTAFVAN	240		
Db	181	YSNLVSGDGGAI	DAKSLTVQGISKL	CVPOBNTAODGCA	QCVVTSFSA	MANEAPTAFIAN	240		
Qy	241	VAGVYRGGGIAA	VQDQGGVSSSTED	PVVSFSRNTAVE	PDGNVARVCGGI	YSYGNVAPL	300		
Db	241	VAGVYRGGGIAA	VQDQGGVSSSTED	PVVSFSRNTAVE	PDGNVARVCGGI	YSYGNVAPL	300		
Qy	301	NNGKTLFLNN	VASPVYIAAKQPT	SGOASNTSNNT	GDGGAIFCQGAQ	-AGSNNSGSVDF	359		
Db	301	NNGKTLFLNN	VASPVYIAEQPT	NGQASNTSDNT	GDGGAIFCQGAQ	AGSNNSGSVDF	360		
Qy	360	GEGVYVFSSN	VAAKGGAIIYAKKL	SVANCGPVQFL	RNIANDGGAIIYLG	ESGELSLSADYG	419		
Db	361	GEGVYVFSSN	VAAKGGAIIYAKKL	SVANCGPVQFL	RNIANDGGAIIYLG	ESGELSLSADYG	420		
Qy	420	DIIFDGNLKR	TAKENADVNGVT	YSSQAI	SMGSGKITTIL	BAKAGHOILFNDP	EMANGN	479	
Db	421	DIIFDGNLKR	TAKENADVNGVT	YSSQAI	SMGSGKITTIL	BAKAGHOILFNDP	EMANGN	480	
Qy	480	NQPAQSSK	LKINDGEGYTGDI	VFANGSSTLYQ	NVTIEQGRIVL	REKAKLSVNSLSQ	TGG	539	
Db	481	NQPAQSSK	LKINDGEGYTGDI	VFANGSSTLYQ	NVTIEQGRIVL	REKAKLSVNSLSQ	TGG	540	
Qy	540	SLYMEAGST	WDFVTPQPPQPP	AAQNIITLSN	LHLSLSLLANNA	VTNPTNP	PPAQDSHP	599	
Db	541	SLYMEAGST	WDFVTPQPPQPP	AAQNIITLSN	LHLSLSLLANNA	VTNPTNP	PPAQDSHP	600	
Qy	600	AVIGSTTAG	SVTISGPIFFED	LDLDTAYDR	YDNLGSKQINVL	KLQGT	KPPANAPSDLT	659	
Db	601	AIIGSTTAG	SVTISGPIFFED	LDLDTAYDR	YDNLGSKQINVL	KLQGT	QPSANAPSDLT	660	
Qy	660	GNEPKYGYQ	GSWKLANDP	TNANGPYTLKAT	WTGTGYNP	GPERSIVASIVP	NSLMSGLI	719	
Db	661	GNEPKYGYQ	GSWKLANDP	TNANGPYTLKAT	WTGTGYNP	GPERSIVASIVP	NSLMSGLI	720	
Qy	720	RSAHSAIQAS	VDGRS	YCRGLW	SVSNFPHY	DRDALGQYRIISG	YSYLGANSYFGSSMF	779	
Db	721	RSAHSAIQAS	VDGRS	YCRGLW	SVSNFPHY	DRDALGQYRIISG	YSYLGANSYFGSSMF	780	
Qy	780	GLATFTEV	GRSKDYVCR	SNHHACIGSV	YLSLTOALCGSY	LYFGDA	PTRASYPGNQHKMT	839	
Db	781	GLATFTEV	GRSKDYVCR	SNHHACIGSV	YLSLTKOALCGSY	LYFGDA	PTRASYPGNQHKMT	840	
Qy	840	SYTPAESD	VRWNNCLAGEI	GAGLP	TVITPSKLYL	NELRPFVQ	AEBSYADHESPT	899	
Db	841	SYTPAESD	VRWNNCLAGEI	GAGLP	TVITPSKLYL	NELRPFVQ	AEBSYADHESPT	900	
Qy	900	QARAFK	SGHLINL	SVPGVGFDR	CSSTHPNKYSF	MAAYICDAY	RTISGTETTL	LSHQBTW	959
Db	901	QARAFK	SGHLINL	SVPGVGFDR	CSSTHPNKYSF	MAAYICDAY	RTISGTETTL	LSHQBTW	960
Qy	960	TTDAFLHAR	GVVVVGR	SMYASL	TSNIEVYGH	GRYEYR	DASRGYGLS	AGSRVRF	1012
Db	961	TTDAFLHAR	GVVVVGR	SMYASL	TSNIEVYGH	GRYEYR	DTSRGYGLS	AGSKVRF	1013

RESULT 6
US-10-498-327-81
; Sequence 81, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761

```
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-81

Query Match          97.4%; Score 5131.5; DB 5; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSPHKPFLSMILAYSCCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60

QY 61 SAGELTLKLNLSIAALPLSCFNLGSLTVLGRGHS�TFENIRTSNGAALSNSAAGL 120
DB 61 SAGELTLKLNLSIAALPLSCFNLGSLTVLGRGHS�TFENIRTSNGAALSNSAAGL 120

QY 121 FTIEGPKELSFNCNLSLLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSFNCNLSLLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNEKPSF 180

QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240

QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240

QY 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300
DB 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300

QY 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359

QY 360 GEGVWFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419
DB 360 GEGVWFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419

QY 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKI TLRKAGHQIILFNDPIEWANGN 479
DB 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKI TLRKAGHQIILFNDPIEWANGN 479

QY 480 NQPAQSSEPLKINDGEGYTDIVFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539
DB 480 NQPAQSSEPLKINDGEGYTDIVFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539

QY 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 599
DB 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 599

QY 541 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 600
DB 541 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 600

; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-81

Query Match          97.4%; Score 5131.5; DB 5; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MQTSPHKPFLSMILAYSCCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60
DB 1 MQTSPHKPFLSMILAYSCCSLNGGVAEIMVPOGIYDGETLTVGPPYTVIGDPSGTTVF 60

QY 61 SAGELTLKLNLSIAALPLSCFNLGSLTVLGRGHS�TFENIRTSNGAALSNSAAGL 120
DB 61 SAGELTLKLNLSIAALPLSCFNLGSLTVLGRGHS�TFENIRTSNGAALSNSAAGL 120

QY 121 FTIEGPKELSFNCNLSLLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSFNCNLSLLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLLNNEKPSF 180

QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240

QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN 240

QY 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300
DB 241 VAGVRGGIAAVQDGGQVSSSTSTEDPVVPSRNTAVFPGNVARVGGGIYSYGNVAF 300

QY 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAQPTSGOASNTSNNGYDGGAI FCKNGAQ - AGSNNSGVSFD 359

QY 360 GEGVWFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419
DB 360 GEGVWFSSNVAAGKGAIAKLSVANCGPVQFLNANDGGAIYLGSGELSLADYG 419

QY 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKI TLRKAGHQIILFNDPIEWANGN 479
DB 420 DIIFDGNLKRKTAKENADVNGVTSSQAI SMGSGGKI TLRKAGHQIILFNDPIEWANGN 479

QY 480 NQPAQSSEPLKINDGEGYTDIVFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539
DB 480 NQPAQSSEPLKINDGEGYTDIVFANGNSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG 539

QY 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 599
DB 540 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 599

QY 541 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 600
DB 541 SLYMEAGSTWDFVTPPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 600
```

[illegible]

RESULT 8

```

US-10-498-327-9
; Sequence 9, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498.327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-9

```

Query Match	97.3%;	Score 5123.5;	DB 5;	Length 1013;
Best Local Similarity	96.6%;	Pred. No. 0;		
Matches 979;	Conservative 22;	Mismatches 11;	Indels 1;	Gaps 1
Qy	1	MQTSFHKFFLSMILAYSCCSLNGGGYAAABIMVPOGIYDGETLTVSFPYTVIGDPSGTTVP	60	
Db	1	MQTSFHKFFLSMILAYSCCSLNGGGYAAABIMVPOGIYDGETLTVSFPYTVIGDPSGTTVP	60	
Qy	61	SAGELTLKLNLDNSIAALPLSCFCGNLLGSAFTVLGRGHSLTFFENIRTSITNGAALSNSAADGL	120	
Db	61	SAGELTLKLNLDNSIAALPLSCFCGNLLGSAFTVLGRGHSLTFFENIRTSITNGAALSNSAAGSL	120	
Qy	121	FTTEGPFKELSFNSCNLSLAVLPAATTKNGSQTPTTTSPSNGTIYSKTDLLLLLNNEKPSF	180	
Db	121	FTTEGPFKELSFNSCNLSLAVLPAATTKNGSQTPTTTSPSNGTIYSKTDLLLLLNNEKPSF	180	
Qy	181	YSNLVSDGGGAIDAKSLTVQGISKLCVFQENITAGDGACQVWTSFSMAWEAPITAFVAN	240	
Db	181	YSNLVSDGGGAIDAKSLTVQGISKLCVFQENITAGDGACQVWTSFSMAWEAPITAFVAN	240	
Qy	241	VAGVRGGGIIAQQDQGGVSSTSTEDPVSFSRNTAVEFDGNVARVGGGITSYSGNVAFI	300	

[illegible]

```

RESULT 9
US-10-498-327-13
; Sequence 13, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-13

```

Query Match 97.3%: Score 5123.5: DB 5: Length 1013:

Best Local Similarity 96.6%; Pred. No. 0;
Matches 979; Conservative 22; Mismatches 11; Indels 1; Gaps 1;
QY 1 MQTSPHKFLLSMILAYSCCLNGGYAAEIMVPOGIYDGETLTVSPFTVIGDPSGTTVP 60
DB 1 MQTSPHKFLLSMILAYSCCLNGGYAAEIMVPOGIYDGETLTVSPFTVIGDPSGTTVP 60
QY 61 SAGELUTLKNLONSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGAALSNSAADGL 120
DB 61 SAGELUTLKNLONSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGAALSNSAADGL 120
QY 121 FTIEGFKELSPNCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLILNNEKFSF 180
DB 121 FTIEGFKELSPNCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLILNNEKFSF 180
QY 181 YSNLVSFGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPFSAMANEAPAFVAN 240
DB 181 YSNLVSFGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPFSAMANEAPAFVAN 240
QY 241 VAGVRGGGIAAVDQGGQVSSSTSTEDPVVSPSRNTAVFPGNVARVGGGIYSYGNVAPL 300
DB 241 VAGVRGGGIAAVDQGGQVSSSTSTEDPVVSPSRNTAVFPGNVARVGGGIYSYGNVAPL 300
QY 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAI FCNGAQ-AGSNNSGSVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAI FCNGAQ-AGSNNSGSVSFD 359
QY 360 GEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLNIAANDGGAIVLGSBGLSLSADYG 419
DB 360 GEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLNIAANDGGAIVLGSBGLSLSADYG 419
QY 420 DIIPDGNLKRXTAKENAADVNGVTVSSQAISMGGGKIITTLRAKAGHOILFNDPIEMANGN 479
DB 420 DIIPDGNLKRXTAKENAADVNGVTVSSQAISMGGGKIITTLRAKAGHOILFNDPIEMANGN 479
QY 480 NQPAQSSSKLLKNDGEGYTDIVFANGSTLVQNTIEQIRIVLREKAKLSVNSLSQTCG 539
DB 480 NQPAQSSSKLLKNDGEGYTDIVFANGSTLVQNTIEQIRIVLREKAKLSVNSLSQTCG 539
QY 540 SLYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSILLANNAVTPNPPAQDSHP 599
DB 540 SLYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSILLANNAVTPNPPAQDSHP 599
QY 599 QARAPKSGHLNLNVPGVGVFDRCSSTHPNKYSFMAAYICDAYRTISGTEITLLSHQETW 959
DB 599 QARAPKSGHLNLNVPGVGVFDRCSSTHPNKYSFMAAYICDAYRTISGTEITLLSHQETW 959
QY 960 TDAFLHARHGVIIRGSGMYASLTNSIEVYGHGRYEYRDSRGSAGSVRF 1012
DB 960 TDAFLHARHGVIIRGSGMYASLTNSIEVYGHGRYEYRDSRGSAGSVRF 1012

US-10-498-327-93
; Sequence 93, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498.327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 93
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-93
Query Match 97.3%; Score 5123.5; DB 5; Length 1013;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 979; Conservative 22; Mismatches 11; Indels 1; Gaps 1;
QY 1 MQTSPHKFLLSMILAYSCCLNGGYAAEIMVPOGIYDGETLTVSPFTVIGDPSGTTVP 60
DB 1 MQTSPHKFLLSMILAYSCCLNGGYAAEIMVPOGIYDGETLTVSPFTVIGDPSGTTVP 60
QY 61 SAGELUTLKNLONSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGAALSNSAADGL 120
DB 61 SAGELUTLKNLONSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNGAALSNSAADGL 120
QY 121 FTIEGFKELSPNCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLILNNEKFSF 180
DB 121 FTIEGFKELSPNCNLSLAVLPAATNKGSOPTTTSTPSNGTIYSKTDLLILNNEKFSF 180
QY 181 YSNLVSFGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPFSAMANEAPAFVAN 240
DB 181 YSNLVSFGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPFSAMANEAPAFVAN 240
QY 241 VAGVRGGGIAAVDQGGQVSSSTSTEDPVVSPSRNTAVFPGNVARVGGGIYSYGNVAPL 300
DB 241 VAGVRGGGIAAVDQGGQVSSSTSTEDPVVSPSRNTAVFPGNVARVGGGIYSYGNVAPL 300
QY 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAI FCNGAQ-AGSNNSGSVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAKOPTSGOASNTSNNGYDGGAI FCNGAQ-AGSNNSGSVSFD 359
QY 360 GEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLNIAANDGGAIVLGSBGLSLSADYG 419
DB 360 GEGVVPFSSNVAAGKGAIAKLSVANGCPVQFLNIAANDGGAIVLGSBGLSLSADYG 419
QY 420 DIIPDGNLKRXTAKENAADVNGVTVSSQAISMGGGKIITTLRAKAGHOILFNDPIEMANGN 479
DB 420 DIIPDGNLKRXTAKENAADVNGVTVSSQAISMGGGKIITTLRAKAGHOILFNDPIEMANGN 479
QY 480 NQPAQSSSKLLKNDGEGYTDIVFANGSTLVQNTIEQIRIVLREKAKLSVNSLSQTCG 539
DB 480 NQPAQSSSKLLKNDGEGYTDIVFANGSTLVQNTIEQIRIVLREKAKLSVNSLSQTCG 539
QY 540 SLYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSILLANNAVTPNPPAQDSHP 599
DB 540 SLYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSILLANNAVTPNPPAQDSHP 599
QY 599 QARAPKSGHLNLNVPGVGVFDRCSSTHPNKYSFMAAYICDAYRTISGTEITLLSHQETW 959
DB 599 QARAPKSGHLNLNVPGVGVFDRCSSTHPNKYSFMAAYICDAYRTISGTEITLLSHQETW 959
QY 960 TDAFLHARHGVIIRGSGMYASLTNSIEVYGHGRYEYRDSRGSAGSVRF 1012
DB 960 TDAFLHARHGVIIRGSGMYASLTNSIEVYGHGRYEYRDSRGSAGSVRF 1012


```
Db 721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGYRIISGYSLGANSYFGSSMF 780
Qy 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 840
Qy 840 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFPSYADHESPTREGD 899
Db 841 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFPSYADHESPTREGD 900
Qy 900 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959
Db 901 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960
Qy 960 TTDAPHLARHGVIVRGSMYASLTSNIEVYHGRYEYRDTSRGYGLSAGSKVRF 1012
Db 961 TTDAPHLARHGVIVRGSMYASLTSNIEVYHGRYEYRDTSRGYGLSAGSKVRF 1013

RESULT 11
US-10-498-327-1
; Sequence 1, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-1
```

```
Query Match 97.2%; Score 5121.5; DB 5; Length 1013;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 978; Conservative 23; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPOGIYDGETLTVSPFYVIGDPSGTTVP 60
Db 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPOGIYDGETLTVSPFYVIGDPSGTTVP 60

Qy 61 SAGELTLKNLNDNSIAALPLSCFNLGSPVVLGRGHSITPENIRTSNNGAALSADGL 120
Db 61 SAGELTLKNLNDNSIAALPLSCFNLGSPVVLGRGHSITPENIRTSNNGAALSADGL 120

Qy 121 FTISGFKELSPNCNLSLAVLPAATTNGSGTPTTTPSPNGTIYSKTDILLANKEFSP 180
Db 121 FTISGFKELSPNCNLSLAVLPAATTNGSGTPTTTPSPNGTIYSKTDILLANKEFSP 180

Qy 181 YSNLVSVDGGGAIDAKSLTVQGISKLCVPEQNTAQDGGACQVTSFSAWANEAPAFVAN 240
Db 181 YSNLVSVDGGGAIDAKSLTVQGISKLCVPEQNTAQDGGACQVTSFSAWANEAPAFVAN 240

Qy 241 VAGVRGGGIAAVQDQGGQVSSSTSTEDPVVFSRNTAVEFGNVARVGGGIIYSVGNVAF 300
Db 241 VAGVRGGGIAAVQDQGGQVSSSTSTEDPVVFSRNTAVEFGNVARVGGGIIYSVGNVAF 300

Qy 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKCKNGAQ-AGSNNGSVSFD 359
Db 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKCKNGAQ-AGSNNGSVSFD 360

Qy 360 GEGVVFSSNVAAGKGGAIYAKGLSVANCGPVQFLRNIAINDGGAIYLGESGELSADYG 419
Db 361 GEGVVFSSNVAAGKGGAIYAKGLSVANCGPVQFLRNIAINDGGAIYLGESGELSADYG 420
```

```
Qy 420 DIIIPDGNLKRRTAKENAADVNGVTVSSQAIISGSGGKITTLRAKACHQILLFNDPIEMANGN 479
Db 421 DIIIPDGNLKRRTAKENAADVNGVTVSSQAIISGSGGKITTLRAKACHQILLFNDPIEMANGN 480
Qy 480 NQPAQSCKLLKINDGEGYTGDIIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQGG 539
Db 481 NQPAQSSEPLKINDGEGYTGDIIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQGG 540
Qy 540 SLYNEAGSTWDPVTPPQPQPPAANQLITLSNLHLSLSLLANNAVTPPTNPPAQDSHP 599
Db 541 SLYNEAGSTLDFVTPPQPQPPAANQLITLSNLHLSLSLLANNAVTPPTNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIFFEDLDLDTAYDRYDNLGSKINVLKQLGTYKPPANAPSIDLTL 659
Db 601 AIIIGSTTAGSVTISGPIFFEDLDLDTAYDRYDNLGSKINVLKQLGTYKPPANAPSIDLTL 660
Qy 660 GNEPMKPYGCGSWKLANDPNTANNPVTTLKATTKTCYVPCGPERVASIVPNSLWGSILDI 719
Db 661 GNEPMKPYGCGSWKLANDPNTANNPVTTLKATTKTCYVPCGPERVASIVPNSLWGSILDI 720
Qy 720 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGYRIISGYSLGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHDDRDLAQGYRIISGYSLGANSYFGSSMF 780
Qy 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 840
Qy 840 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFPSYADHESPTREGD 899
Db 841 SYTPABESDVRWNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAFPSYADHESPTREGD 900
Qy 900 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959
Db 901 QARAFKSGHLNLSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960
Qy 960 TTDAPHLARHGVIVRGSMYASLTSNIEVYHGRYEYRDTSRGYGLSAGSKVRF 1012
Db 961 TTDAPHLARHGVIVRGSMYASLTSNIEVYHGRYEYRDTSRGYGLSAGSKVRF 1013

RESULT 12
US-10-498-327-3
; Sequence 3, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCE: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-498-327-3
```

```
Query Match 97.2%; Score 5121.5; DB 5; Length 1013;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 978; Conservative 23; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPOGIYDGETLTVSPFYVIGDPSGTTVP 60
Db 1 MQTSFHKFFLSMILAYSCCSLGGGYAAEIMVPOGIYDGETLTVSPFYVIGDPSGTTVP 60

Qy 61 SAGELTLKNLNDNSIAALPLSCFNLGSPVVLGRGHSITPENIRTSNNGAALSADGL 120
Db 61 SAGELTLKNLNDNSIAALPLSCFNLGSPVVLGRGHSITPENIRTSNNGAALSADGL 120
```

```
QY 121 FTIEGPKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180
DB 121 FTIEGPKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNEKPSF 180
QY 181 YSNLSVGGGGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPFSAMANEAPAFVAN 240
DB 181 YSNLSVGGGGAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPFSAMANEAPAFVAN 240
QY 241 VAGVGGGGAIAVQDGGQGVSSSTSTEDPVVPSRNTAVEFDGNVAVRGGGIYSYGNVAPL 300
DB 241 VAGVGGGGAIAVQDGGQGVSSSTSTEDPVVPSRNTAVEFDGNVAVRGGGIYSYGNVAPL 300
QY 301 NNGKTLFLNNVAPVIAAKQPTSGQASNTSNNGYDGGGAIKCNKGAQ-AGSNNSGSVSD 359
DB 301 NNGKTLFLNNVAPVIAAKQPTSGQASNTSNNGYDGGGAIKCNKGAQ-AGSNNSGSVSD 359
QY 360 GEGVVFSSNVAAGKGGAIAKLSVANGCPVQFLRNANDGGAIYLGESGELSLSADYG 419
DB 360 GEGVVFSSNVAAGKGGAIAKLSVANGCPVQFLRNANDGGAIYLGESGELSLSADYG 419
QY 420 DIIFDGNLARTAKENAADVNGVTSSQAI SMGSGGKITTLLRAKAGHQILFNDPIEMANG 479
DB 420 DIIFDGNLARTAKENAADVNGVTSSQAI SMGSGGKITTLLRAKAGHQILFNDPIEMANG 479
QY 480 NQPAQSSKLLKINDGEGYTGDIVPANGSSTLYONVTIEQGRIVLREKAKLSVNSLSQTCG 539
DB 480 NQPAQSSKLLKINDGEGYTGDIVPANGSSTLYONVTIEQGRIVLREKAKLSVNSLSQTCG 539
QY 540 SLYMEAGSTWDTVTPPOPPQAPNOLITLNLHLSLSLLANNATVNPPTPPAQDSHP 599
DB 540 SLYMEAGSTWDTVTPPOPPQAPNOLITLNLHLSLSLLANNATVNPPTPPAQDSHP 599
QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKIDVLKQLGTQPSANAPSDTL 659
DB 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKIDVLKQLGTQPSANAPSDTL 659
QY 660 GNEMPKYQGSWKLAWDNTANGPYTLKATWTKGTGNPGPERRVASLVPNSLWGSILDI 719
DB 660 GNEMPKYQGSWKLAWDNTANGPYTLKATWTKGTGNPGPERRVASLVPNSLWGSILDI 719
QY 720 RSHSAIAQSVGGRSVCRLWVSGVSNFFVHDDALGQGVYISGGYSLGANSYFGSSMF 779
DB 720 RSHSAIAQSVGGRSVCRLWVSGVSNFFVHDDALGQGVYISGGYSLGANSYFGSSMF 779
QY 780 GLAFTVFGRSKDYVCRSNHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 839
DB 780 GLAFTVFGRSKDYVCRSNHACIGSVYLSSTQALCGSYLFGDAFIRASYGFGNQHMKT 839
QY 840 SYTFABESDVRWNNCLAGIAGLPIVITPSSKLYLNELRPFVQABPSYADHESFTEGD 899
DB 840 SYTFABESDVRWNNCLAGIAGLPIVITPSSKLYLNELRPFVQABPSYADHESFTEGD 899
QY 900 QABAPKSGHLLNLVSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTTSGTETLLSHQETW 959
DB 900 QABAPKSGHLLNLVSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTTSGTETLLSHQETW 959
QY 960 TTDAPFLARHGVIIVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSVRF 1012
DB 960 TTDAPFLARHGVIIVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSVRF 1012
```

RESULT 13

```
US-10-701-844-43
; Sequence 43, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10701.844
; CURRENT FILING DATE: 2003-11-04
```

```
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-43

Query Match 97.1%; Score 5116; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGGIYDGETLTVSPFYTVIGDPSTTVPSAGELTLKNLDNSIALPLSCFNGLLGS 88
DB 1 EIMVPGGIYDGETLTVSPFYTVIGDPSTTVPSAGELTLKNLDNSIALPLSCFNGLLGS 60
QY 89 FTVLGRGHSHTPENIRTSNGAALSNSAADGLFTIEGPKELSFNSCNSLLAVLPAATTNK 148
DB 61 FTVLGRGHSHTPENIRTSNGAALSNSAADGLFTIEGPKELSFNSCNSLLAVLPAATTNK 120
QY 149 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVP 208
DB 121 GSQTPTTSTPSNGTIYSKTDLLLNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVP 180
QY 209 QNTAQAAGGACQVVTSPFSAMANEAPAFVANVAVRGGGIAAIVQDGGQGVSSSTEDP 268
DB 181 QNTAQAAGGACQVVTSPFSAMANEAPAFVANVAVRGGGIAAIVQDGGQGVSSSTEDP 240
QY 269 VVPSRNTAVEFDGNVAVRGGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQAS 328
DB 241 VVPSRNTAVEFDGNVAVRGGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQAS 300
QY 329 NTSNNYDGGGAIFCKNGAQAQSNNSGSVSFDGEGVVFSSNVAAGKGGAIAKLSVANC 388
DB 301 NTSNNYDGGGAIFCKNGAQAQSNNSGSVSFDGEGVVFSSNVAAGKGGAIAKLSVANC 360
QY 389 GPVQFLRNANDGGAIYLGESGELSLSADYGDIIIFDGNLKRKTAKENAADVNGVTSSQAI 448
DB 361 GPVQFLRNANDGGAIYLGESGELSLSADYGDIIIFDGNLKRKTAKENAADVNGVTSSQAI 420
QY 449 SMGSGGKIITTLRAKAGHQILFNDPIEMANGNQPAOSSKLLKINDGEGYTGDIVPANGSS 508
DB 421 SMGSGGKIITTLRAKAGHQILFNDPIEMANGNQPAOSSKLLKINDGEGYTGDIVPANGSS 480
QY 509 TLYQNTVIEQGRIVLREKAKLSVNSLSQTCGSLYMEAGSTWDTVTPPQPPPAANQLIT 568
DB 481 TLYQNTVIEQGRIVLREKAKLSVNSLSQTCGSLYMEAGSTWDTVTPPQPPPAANQLIT 540
QY 569 LSNLHLSLSLLANNATVNPPTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDR 628
DB 541 LSNLHLSLSLLANNATVNPPTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDR 600
QY 629 YDWLGSNQKINVLKQLGTQKPPANAPSDTLTGNEPKYQGSWKLAWDNTANGPYTL 688
DB 601 YDWLGSNQKINVLKQLGTQKPPANAPSDTLTGNEPKYQGSWKLAWDNTANGPYTL 660
QY 689 KATWTKGTGNPGPERRVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLWVSGVSNFF 748
DB 661 KATWTKGTGNPGPERRVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLWVSGVSNFF 720
QY 749 YHDDRALGQGVYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVYVCRSNHACIGSVY 808
DB 721 YHDDRALGQGVYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVYVCRSNHACIGSVY 780
QY 809 LSTQQAALCGSYLFGDAFIRASYGFGNQHMKTSTFAEESDVRWNNCLAGEICAGIPIVI 868
DB 781 LSTQQAALCGSYLFGDAFIRASYGFGNQHMKTSTFAEESDVRWNNCLAGEICAGIPIVI 840
QY 869 TPSKLYLNELRPFVQABPSYADHESFTEGDQAPAPKSGHLLNLVSPVGVKFCRCSSTHP 928
```

Db 841 TPSKLYLNELRPVQAEFSYADHESFTTEGQARAFKSGHLNLSVPVGVKFDRCSSHP 900
Qy 929 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTNSIEVY 988
Db 901 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTNSIEVY 960
Qy 989 GHGREYRDASRGYGLSAGSRVP 1012
Db 961 GHGREYRDASRGYGLSAGSRVP 984

RESULT 14
US-10-766-711-43
; Sequence 43, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766.711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-43

Query Match 97.1%; Score 5116; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFSGAGELTLKNDNSIAALPLSCFNLG 88
Db 1 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFSGAGELTLKNDNSIAALPLSCFNLG 60

Qy 89 FTVLGRGHSITPENIRSTNGAALSNSAADGLFTIEGFKELSFNCNSLLAVLPAATNK 148
Db 61 FTVLGRGHSITPENIRSTNGAALSNSAADGLFTIEGFKELSFNCNSLLAVLPAATNK 120

Qy 149 GSQPTTTTSPNGTIVSKTDLNNEKFSFYGNLVSGDGAIDAKSLTVQGISKLCVP 208
Db 121 GSQPTTTTSPNGTIVSKTDLNNEKFSFYGNLVSGDGAIDAKSLTVQGISKLCVP 180

Qy 209 QENTAQADGGACQVTVSFSAMANEAPAFVANAVGRGGIAAVDQGGVSSSTSTEDP 268
Db 181 QENTAQADGGACQVTVSFSAMANEAPAFVANAVGRGGIAAVDQGGVSSSTSTEDP 240

Qy 269 VVSFRTAVFEDGNVAVRGGIYSYGNVAFNNKTLFLNNVASPVVIAAKOPTSGAS 328
Db 241 VVSFRTAVFEDGNVAVRGGIYSYGNVAFNNKTLFLNNVASPVVIAAKOPTSGAS 300

Qy 329 NTSNNYGDGGAIFCKNGAQSNNSGSVSPDGEVWFFSSNVAAGKGAIIYAKKLSVANC 388
Db 301 NTSNNYGDGGAIFCKNGAQSNNSGSVSPDGEVWFFSSNVAAGKGAIIYAKKLSVANC 360

Qy 389 GPVQFLNIANDGAIYLGESGELSLSDYGDIIIPDGNLKRTAKENAADVNGVTVSSQAI 448
Db 361 GPVQFLNIANDGAIYLGESGELSLSDYGDIIIPDGNLKRTAKENAADVNGVTVSSQAI 420

Qy 449 SMGSGGKITTLURAKAGHQILFNDPIEMANGNQPAQSSKLLKINDGSGYTGDIYFANGSS 508
Db 421 SMGSGGKITTLURAKAGHQILFNDPIEMANGNQPAQSSKLLKINDGSGYTGDIYFANGSS 480

Qy 509 TLYQNVITBQRIIVLREKAKLSVNSLSQTGSLTMEAGSTWDFVTPPQPPQPPAANQLIT 568
Db 481 TLYQNVITBQRIIVLREKAKLSVNSLSQTGSLTMEAGSTWDFVTPPQPPQPPAANQLIT 540

Qy 569 LSNLHLSLSLLANNVNTNPPAODSHPAVIGSTTAGSVTISGPIFPEDLDDTAYDR 628
Db 541 LSNLHLSLSLLANNVNTNPPAODSHPAVIGSTTAGSVTISGPIFPEDLDDTAYDR 600

Qy 629 YDMLGSNOKNLVNLKQLGTTPPANAPSDLTGLNEMPKYQGGSKLAWDPNTANNGPYTL 688
Db 601 YDMLGSNOKNLVNLKQLGTTPPANAPSDLTGLNEMPKYQGGSKLAWDPNTANNGPYTL 660

Qy 689 KATWTKGTYNPGPVERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNFF 748
Db 661 KATWTKGTYNPGPVERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLMWGVSNNFF 720

Qy 749 YHDRDALGQGYRYISGGYSLGANSYFGSSNPFGLAFTVFGRSKDYVYCRSNHHAICISVY 808
Db 721 YHDRDALGQGYRYISGGYSLGANSYFGSSNPFGLAFTVFGRSKDYVYCRSNHHAICISVY 780

Qy 809 LSTQOALCGSYLPQDAFIRASYGFGNOMKTSYTFAESDVRWNNCLAGEIGAGLPIVI 868
Db 781 LSTQOALCGSYLPQDAFIRASYGFGNOMKTSYTFAESDVRWNNCLAGEIGAGLPIVI 840

Qy 869 TPSKLYLNELRPVQAEFSYADHESFTTEGQARAFKSGHLNLSVPVGVKFDRCSSHP 928
Db 841 TPSKLYLNELRPVQAEFSYADHESFTTEGQARAFKSGHLNLSVPVGVKFDRCSSHP 900

Qy 929 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTNSIEVY 988
Db 901 NKYSFMAAYICDAYRTISGTETLLSHQETWTTDAFLARHGTVVVRGSMYASLTNSIEVY 960

Qy 989 GHGREYRDASRGYGLSAGSRVP 1012
Db 961 GHGREYRDASRGYGLSAGSRVP 984

RESULT 15
US-10-931-779-43
; Sequence 43, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BP104
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 43
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-43

Query Match 97.1%; Score 5116; DB 5; Length 984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFSGAGELTLKNDNSIAALPLSCFNLG 88
Db 1 EIMVPOGIYDGETLTVSPFYTVIGDPGQTVFSGAGELTLKNDNSIAALPLSCFNLG 60

Qy 89 FTVLGRGHSITPENIRSTNGAALSNSAADGLFTIEGFKELSFNCNSLLAVLPAATNK 148
Db 61 FTVLGRGHSITPENIRSTNGAALSNSAADGLFTIEGFKELSFNCNSLLAVLPAATNK 120

Qy 149 GSQPTTTTSPNGTIVSKTDLNNEKFSFYGNLVSGDGAIDAKSLTVQGISKLCVP 208
Db 121 GSQPTTTTSPNGTIVSKTDLNNEKFSFYGNLVSGDGAIDAKSLTVQGISKLCVP 180

Qy 209 QENTAQADGGACQVTVSFSAMANEAPAFVANAVGRGGIAAVDQGGVSSSTSTEDP 268
Db 181 QENTAQADGGACQVTVSFSAMANEAPAFVANAVGRGGIAAVDQGGVSSSTSTEDP 240

269 VVSFSPNTAVBPDGNVAVGGGIYSGNVAFPLNNGKTLPLANNVSPVVIKAKOPTSGOAS 328
Db VVSFSPNTAVBPDGNVAVGGGIYSGNVAFPLNNGKTLPLANNVSPVVIKAKOPTSGOAS 300
329 NTSNNYGGGGAIFCKNGAQAAGSNSSGVSFDGEGVVPFSSNVAAAGGGAIAKAKLSVANC 388
Db NTSNNYGGGGAIFCKNGAQAAGSNSSGVSFDGEGVVPFSSNVAAAGGGAIAKAKLSVANC 360
389 GPVQFLRNANDGGAIIYGESGELSADYDGIIFDGNLKRTAKNAADVNGVTVSSQAI 448
Db GPVQFLRNANDGGAIIYGESGELSADYDGIIFDGNLKRTAKNAADVNGVTVSSQAI 420
449 SNGSGGKIITLAKAGHOITLNDPILFEMANGNNOQAQSSKLLKINDGEGVTGDIIVANGSS 508
Db SNGSGGKIITLAKAGHOITLNDPILFEMANGNNOQAQSSKLLKINDGEGVTGDIIVANGSS 480
509 TLQYQNTTIEQRIVLREKAKLSVNSLSQGGSLYMEAGSTWDFVTPQPPQPPAANQLIT 568
Db TLQYQNTTIEQRIVLREKAKLSVNSLSQGGSLYMEAGSTWDFVTPQPPQPPAANQLIT 540
569 LSNLHLSLSLLANNAVTPNPPAQDSHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDR 628
Db LSNLHLSLSLLANNAVTPNPPAQDSHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDR 600
629 YDWLGSNQKINVLKQLGTHKPPANAPSDLTILGNEMPKYQYQSGWKLAWDPNTANNPYTL 688
Db YDWLGSNQKINVLKQLGTHKPPANAPSDLTILGNEMPKYQYQSGWKLAWDPNTANNPYTL 660
689 KATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFF 748
Db KATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNFF 720
749 YHDRDALGQYRIYISGGYSLGANSYFGSSMFLGAPTEVFGRSKDYVVCRSNHACIGSVY 808
Db YHDRDALGQYRIYISGGYSLGANSYFGSSMFLGAPTEVFGRSKDYVVCRSNHACIGSVY 780
809 LSTQQAALCSYLFQDAFIASVFGNQHKMTSTTPAESDVRDNNCLAGEIGAGLPYVI 868
Db LSTQQAALCSYLFQDAFIASVFGNQHKMTSTTPAESDVRDNNCLAGEIGAGLPYVI 840
869 TPKLYNLRLPPVQAEPFVSADHESFTEEGDQARAFKSGHLLNLSVPVGVKFCRCSSTHP 928
Db TPKLYNLRLPPVQAEPFVSADHESFTEEGDQARAFKSGHLLNLSVPVGVKFCRCSSTHP 900
929 NKYSFMAAYICDAYRTISGTETLLSHQETWTTTDAFLARHGTVVVRGSMYASLTSNIEVY 988
Db NKYSFMAAYICDAYRTISGTETLLSHQETWTTTDAFLARHGTVVVRGSMYASLTSNIEVY 960
989 GHGRYERDASRGYGLSAGSRVRF 1012
Db GHGRYERDASRGYGLSAGSRVRF 984

RESULT 16
US-10-498-327-17
; Sequence 17, Application US/10498327
; Publication No. US20050106162A1
; GENERAL INFORMATION:
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
; FILE REFERENCES: 002441.00085
; CURRENT APPLICATION NUMBER: US/10/498,327
; CURRENT FILING DATE: 2004-06-10
; PRIOR APPLICATION NUMBER: PCT/IB02/05761
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 262
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis

US-10-498-327-17

Query Match 97.1%; Score 5112.5; DB 5; Length 1013;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 977; Conservative 23; Mismatches 12; Indels 1; Gaps 1;
QY 1 MQTSFHKFPLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPSGTTVF 60
Db 1 MQTSFHKFPLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFVTVIGDPSGTTVF 60
61 SAGELTLKVLNDNSTAALPLSCFGLNLSGFTVLGRHSITFEENIRTSITNGAALSNSAADGL 120
Db 61 SAGELTLKVLNDNSTAALPLSCFGLNLSGFTVLGRHSITFEENIRTSITNGAALSNSAADGL 120
121 FTIEGFKELSFNSCNSLLAVLPAATNKGSTPTTTSTPSNGTIYKTDLLLNNEKPSF 180
Db 121 FTIEGFKELSFNSCNSLLAVLPAATNKGSTPTTTSTPSNGTIYKTDLLLNNEKPSF 180
181 YSNLVSGDGGGAIDAKSLTVQGISKL CVFOENTAOADGGACQVVTFSFMANEAPAFVAN 240
Db 181 YSNLVSGDGGGAIDAKSLTVQGISKL CVFOENTAOADGGACQVVTFSFMANEAPAFVAN 240
241 VAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSFRNTAVEFDGNVARVGGGIYSGNVAF 300
Db 241 VAGVRGGGIAAVQDQOQGVSSSTSTEDPVVSFRNTAVEFDGNVARVGGGIYSGNVAF 300
301 NNGKTLPLNNVASPVYIAAKOPTSGOASNTSNYGDGGAIFCKNGAQ-AGSNNSVSFD 359
Db 301 NNGKTLPLNNVASPVYIAAKOPTSGOASNTSNYGDGGAIFCKNGAQAGSNNSVSFD 360
360 GEGVVPFSSNVAAAGKGGAIYAKKLSVANC GPVQFLNNDGGAIIYLGSGSLSADYG 420
Db 360 GEGVVPFSSNVAAAGKGGAIYAKKLSVANC GPVQFLNNDGGAIIYLGSGSLSADYG 420
420 DIIFDGNLKRTAKENAAVNGVTVSSQAI SMGSGGKIITLAKAGHOITLNDPDIEMANGN 479
Db 420 DIIFDGNLKRTAKENAAVNGVTVSSQAI SMGSGGKIITLAKAGHOITLNDPDIEMANGN 480
480 NPOASSKLLKINDGEGYTGDIIVFANGNSTLQNTVIEQRIVLREKAKLSVNSLSQGG 539
Db 480 NPOASSKLLKINDGEGYTGDIIVFANGNSTLQNTVIEQRIVLREKAKLSVNSLSQGG 540
540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599
Db 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600
600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKINVLKQLGTYPPANAPSDLT 659
Db 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKIDVLKQLGTYPPANAPSDLT 660
660 GNEMPKYQYQSGWKLAWDPNTANNPYTLKATWTKTGNPGERVASLVPNSLWGSILDI 719
Db 660 GNEMPKYQYQSGWKLAWDPNTANNPYTLKATWTKTGNPGERVASLVPNSLWGSILDI 720
720 RSHASAIQASVDGRSYCRGLWVSGVSNFFYHDDALGQGYRIYSGYSLGANSYFGSSMF 779
Db 720 RSHASAIQASVDGRSYCRGLWVSGVSNFFYHDDALGQGYRIYSGYSLGANSYFGSSMF 780
780 GLAFTEVFGRSKDYVVCRSNHACIGSVYLSQQAALCSYLFQDAFIASVFGNQHKMT 839
Db 780 GLAFTEVFGRSKDYVVCRSNHACIGSVYLSQQAALCSYLFQDAFIASVFGNQHKMT 840
840 SYTFAESDVRDNNCLAGEIGAGLPYITPSKLYNLRLPPVQAEPFVSADHESFTEEGD 899
Db 840 SYTFAESDVRDNNCLAGEIGAGLPYITPSKLYNLRLPPVQAEPFVSADHESFTEEGD 900
900 QARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959
Db 900 QARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960
960 TTDAFLARHGTVVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
Db 960 TTDAFLARHGTVVVRGSMYASLTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1013

```
RESULT 17
US-10-701-844-15
; Sequence 15, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-15

Query Match      97.0%; Score 5106.5; DB 4; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

Qy      1  MQTSPHKPFLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFVTVIGDPSGTTVP 60
Db      1  MQTSPHKPFLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFVTVIGDPSGTTVP 60

Qy      61  SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITFENIRTSNGAALSNSAADGL 120
Db      61  SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITFENIRTSNGAALSNSAADGL 120

Qy      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180
Db      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180

Qy      181  YSNLVSDGGGAIKAKLSVQISKLQVQENTADGGACQVTSFSAANEAPAFVAN 240
Db      181  YSNLVSDGGGAIKAKLSVQISKLQVQENTADGGACQVTSFSAANEAPAFVAN 240

Qy      241  VAGVGGGIAAQQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAF 300
Db      241  VAGVGGGIAAQQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAF 300

Qy      301  NNGKTLFLNNVASFVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAGSNNSGVSFD 360
Db      301  NNGKTLFLNNVASFVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAGSNNSGVSFD 360

Qy      360  GEGVVFPSSNAAGGGAIAKLSVANCVPQVFLRNTANDGGAIYLGESGELSADYG 419
Db      361  GEGVVFPSSNAAGGGAIAKLSVANCVPQVLLGNANTANDGGAIYLGESGELSADYG 420

Qy      420  DIIFDGNLKRITAKENAADVNGVTYSSQAISSMGSGKITTLLAKAGHQLTFNDPIEMANG 479
Db      421  DMIFDGNLKRITAKENAADVNGVTYSSQAISSMGSGKITTLLAKAGHQLTFNDPIEMANG 480

Qy      480  NQPAQSSKLLKINDGEGVTGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG 539
Db      481  NQPAQSSKLLKINDGEGVTGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGG 540

Qy      540  SLYNEAGSTMDPVPQPPQAPAAQLITLNLHLISLSSLLANNAVTPNPPNPAQDSHP 599
Db      541  SLYNEAGSTMDPVPQPPQAPAAQLITLNLHLISLSSLLANNAVTPNPPNPAQDSHP 600

Qy      600  AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKINYLXLQLGTPKPPANAPSIDL 659
Db      601  AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKIDVLKLQLGTQPPANAPSIDL 660
```

```
Qy      660  GNEMPKYGYGSKWLANDPNTANNPPTLAKATWTKTGYNCPERVASILVPSNLGSIILDI 719
Db      661  GNEMPKYGYGSKWLANDPNTANNPPTLAKATWTKTGYNCPERVASILVPSNLGSIILDI 720

Qy      720  RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHRRDALQGGYRIISGGYSLGANSYFGSSMF 779
Db      721  RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHRRDALQGGYRIISGGYSLGANSYFGSSMF 780

Qy      780  GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTQOALCGSYLFGDAFIRASYGFGNQHMKT 839
Db      781  GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSSTQOALCGSYVFGDAFIRASYGFGNQHMKT 840

Qy      840  SYTFAESDVRWNNCLAGHIGAGLPVITPSPKLYINELAPPVQOAFSYADHESFTEBGD 899
Db      841  SYTFAESDVRWNNCLAGHIGAGLPVITPSPKLYINELAPPVQOAFSYADHESFTEBGD 900

Qy      900  QARAPKSHLNLISVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETTLSSHORTW 959
Db      901  QARAPKSHLNLISVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETTLSSHORTW 960

Qy      960  TTDAPHLARHGVVVRGSMYASLTNSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
Db      961  TTDAPHLARHGVVVRGSMYASLTNSNIEVYGHGRYERDASRGYGLSAGSKVRF 1013
```

```
RESULT 18
US-10-701-844-16
; Sequence 16, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-16
```

```
Query Match      97.0%; Score 5106.5; DB 4; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

Qy      1  MQTSPHKPFLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFVTVIGDPSGTTVP 60
Db      1  MQTSPHKPFLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFVTVIGDPSGTTVP 60

Qy      61  SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITFENIRTSNGAALSNSAADGL 120
Db      61  SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITFENIRTSNGAALSNSAADGL 120

Qy      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180
Db      121  FTIEGPKELSFNSCNLSLAVLPAATNKGSTPTTTPSPNGTIYSKTDLLLLNNEKPSF 180

Qy      181  YSNLVSDGGGAIKAKLSVQISKLQVQENTADGGACQVTSFSAANEAPAFVAN 240
Db      181  YSNLVSDGGGAIKAKLSVQISKLQVQENTADGGACQVTSFSAANEAPAFVAN 240

Qy      241  VAGVGGGIAAQQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAF 300
Db      241  VAGVGGGIAAQQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARGGGIYSYGNVAF 300

Qy      301  NNGKTLFLNNVASFVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAGSNNSGVSFD 359
```


Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAI FCKNGAQAAGSNNSGVSF 360
Qy 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGA IYLGESGELSADYG 419
Db 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGA IYLGESGELSADYG 420
Qy 420 DIIPDGNLKRRTAKENADVGVTVSSQAI SMGSGGKIITLRAKAGHOILFNDPIEMANGN 479
Db 421 DIIPDGNLKRRTAKENADVGVTVSSQAI SMGSGGKIITLRAKAGHOILFNDPIEMANGN 480
Qy 480 NQPAOSSKLLKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 539
Db 481 NQPAOSSSEPLKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 540
Qy 540 SLYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSLLANNVTPNPPPAQDSHP 599
Db 541 SLYMEAGSTLDFVTPPOPPQPPAANQLITLSNLHLSLSLLANNVTPNPPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIPFEDIDDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDLT 659
Db 601 AVIGSTTAGSVTISGPIPFEDIDDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDLT 660
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDI 719
Db 661 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDI 720
Qy 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLALGQGYRYISGGYSLGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLALGQGYRYISGGYSLGANSYFGSSMF 780
Qy 780 GLAFTVFGRSKDYVYVCRSNHACIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 839
Db 781 GLAFTVFGRSKDYVYVCRSNHACIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 840
Qy 840 SYTFABESDVRDNNCLAGIAGLPIVITPSPKLYLNLRLPFPVQABFSYADHESFTEEG 899
Db 841 SYTFABESDVRDNNCLAGIAGLPIVITPSPKLYLNLRLPFPVQABFSYADHESFTEEG 900
Qy 900 QARAFKSGHLLNLSPVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 959
Db 901 QARAFKSGHLLNLSPVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 960
Qy 960 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGYEYRDASRGYGLSAGSVRFP 1012
Db 961 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGYEYRDASRGYGLSAGSVRFP 1013

RESULT 19
US-10-766-711-15
; Sequence 15, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-15

Query Match 97.0%; Score 5106.5; DB 4; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

Qy 1 NQTSFHKFFLNLAYSCCSLNGGGYAAABIMVPOGIYDGETLTYSFPYTVIGDPSTGTVF 60
Db 1 NQTSFHKFFLNLAYSCCSLNGGGYAAABIMVPOGIYDGETLTYSFPYTVIGDPSTGTVF 60
Qy 61 SAGELTUKNLONSIAALPLSCFNLGLSFTVLGRHSITTFENIRISTNGAALSNAADGL 120
Db 61 SAGELTUKNLONSIAALPLSCFNLGLSFTVLGRHSITTFENIRISTNGAALSNAADGL 120
Qy 121 FTIEGFKELSPNCSNLSLAVLPAATTNKSGSQPTPTTTPSNGTIYKSTDLLLNNKEKFSF 180
Db 121 FTIEGFKELSPNCSNLSLAVLPAATTNKSGSQPTPTTTPSNGTIYKSTDLLLNNKEKFSF 180
Qy 181 YSNLVSGDGAIDAKSLTVQGISLKVQENTAQADGACQVVTSPSAMANEAPAFVAN 240
Db 181 YSNLVSGDGAIDAKSLTVQGISLKVQENTAQADGACQVVTSPSAMANEAPAFVAN 240
Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEFDGNVAVGVGGIYSYGNVAF 300
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVEFDGNVAVGVGGIYSYGNVAF 300
Qy 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAI FCKNGAQAAGSNNSGVSF 359
Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAI FCKNGAQAAGSNNSGVSF 360
Qy 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGA IYLGESGELSADYG 419
Db 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVOFLNTANDGGA IYLGESGELSADYG 420
Qy 420 DIIPDGNLKRRTAKENADVGVTVSSQAI SMGSGGKIITLRAKAGHOILFNDPIEMANGN 479
Db 421 DIIPDGNLKRRTAKENADVGVTVSSQAI SMGSGGKIITLRAKAGHOILFNDPIEMANGN 480
Qy 480 NQPAOSSKLLKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 539
Db 481 NQPAOSSSEPLKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTCG 540
Qy 540 SLYMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSLLANNVTPNPPPAQDSHP 599
Db 541 SLYMEAGSTLDFVTPPOPPQPPAANQLITLSNLHLSLSLLANNVTPNPPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIPFEDIDDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDLT 659
Db 601 AVIGSTTAGSVTISGPIPFEDIDDDTAYDRYDMLGSKINVLKQLGTQPPANAPSDLT 660
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDI 719
Db 661 GNEMPKYGYQGSWKLAWDPNTANNNGPYTLKATWTKGTGYNPGERVASLVPNSLWGSILDI 720
Qy 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLALGQGYRYISGGYSLGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLALGQGYRYISGGYSLGANSYFGSSMF 780
Qy 780 GLAFTVFGRSKDYVYVCRSNHACIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 839
Db 781 GLAFTVFGRSKDYVYVCRSNHACIGSVLSTQOALCGSVLFGDAFIRASYGFGNOHMT 840
Qy 840 SYTFABESDVRDNNCLAGIAGLPIVITPSPKLYLNLRLPFPVQABFSYADHESFTEEG 899
Db 841 SYTFABESDVRDNNCLAGIAGLPIVITPSPKLYLNLRLPFPVQABFSYADHESFTEEG 900
Qy 900 QARAFKSGHLLNLSPVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 959
Db 901 QARAFKSGHLLNLSPVPGVKFDRCSSTHPNKYSFMAAYICDAVRTISGTTLLSHQETW 960
Qy 960 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGYEYRDASRGYGLSAGSVRFP 1012
Db 961 TTDAPFLARHGVIVRGSMYASLTNSIEVYGHGYEYRDASRGYGLSAGSVRFP 1013

RESULT 20
US-10-766-711-16
; Sequence 16, Application US/10766711
; Publication No. US20040137005A1

; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-16

Query Match 97.0%; Score 5106.5; DB 4; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

Qy 1 MQTSPHKFPLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF 60
Db 1 MQTSPHKFPLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF 60

Qy 61 SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITPENIRTSNGAALSADGL 120
Db 61 SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITPENIRTSNGAALSADGL 120

Qy 121 FTIEGFKELSPNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180
Db 121 FTIEGFKELSPNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180

Qy 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTAQDGGACQVTFSPSAMANEAPAFVAN 240
Db 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTAQDGGACQVTFSPSAMANEAPAFVAN 240

Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAPL 300
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAPL 300

Qy 301 NNGKTLFLNNVASPVYIAAKQPTSOASNTSNYGDGGAIFCKNGAQ-AGSNNGSVSFD 359
Db 301 NNGKTLFLNNVASPVYIAAKQPTSOASNTSNYGDGGAIFCKNGAQ-AGSNNGSVSFD 359

Qy 360 GEGVVFSSNVAAGKGAIIYAKKLSVANCGPVQFLGNANDGGAIIYLGESGELSADYG 419
Db 360 GEGVVFSSNVAAGKGAIIYAKKLSVANCGPVQFLGNANDGGAIIYLGESGELSADYG 419

Qy 420 DIIFDGNLKRKTAKENAADVNGTVSSQAIISWGSGGKIITTLAKAGHQLFPNDPIEMANGN 479
Db 420 DIIFDGNLKRKTAKENAADVNGTVSSQAIISWGSGGKIITTLAKAGHQLFPNDPIEMANGN 479

Qy 480 NQPAQSSKLLKINDGEGVTGDIIVFANGSSTLYQNTVIEQGRIVLRKAKLSVNSISQGG 539
Db 480 NQPAQSSKLLKINDGEGVTGDIIVFANGSSTLYQNTVIEQGRIVLRKAKLSVNSISQGG 539

Qy 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLNLHLSSLANNANVTNPPNPPAQDSHP 599
Db 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLNLHLSSLANNANVTNPPNPPAQDSHP 599

Qy 600 AVIGSTTAGSVTIISGPIFFEDLDDTAYDRYDNLGSKINVLKQLGTPPANAPSDTL 659
Db 600 AVIGSTTAGSVTIISGPIFFEDLDDTAYDRYDNLGSKINVLKQLGTPPANAPSDTL 659

Qy 660 GNEAPKYGQSGWKLWDPNTANNPPTLKATWTGTGYNPCGPVRSIVLPSNLGSIIDI 719
Db 660 GNEAPKYGQSGWKLWDPNTANNPPTLKATWTGTGYNPCGPVRSIVLPSNLGSIIDI 719

Qy 720 RSAHSIAIQASVDGRSICRGLWVSGVSNFFYHDDRALGQGYRIISGGYSLGANSYFGSSMF 779
Db 720 RSAHSIAIQASVDGRSICRGLWVSGVSNFFYHDDRALGQGYRIISGGYSLGANSYFGSSMF 779

Qy 780 GLATEVFGRSKDYVVCRSNHACIGSVYLSQOALCGSYLFGDAFIRASVFGNGHMKT 839
Db 780 GLATEVFGRSKDYVVCRSNHACIGSVYLSQOALCGSYLFGDAFIRASVFGNGHMKT 839

Qy 840 SYTFAESDVRWDDNNCLAGIEGAGLPIVITPSKLYLNELRPFVQAEPSYADHESPTSEGD 899
Db 840 SYTFAESDVRWDDNNCLAGIEGAGLPIVITPSKLYLNELRPFVQAEPSYADHESPTSEGD 899

Qy 900 QARAPKSGHLLNLSPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTTTLLSHQBTW 959
Db 900 QARAPKSGHLLNLSPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTTTLLSHQBTW 959

Qy 960 TTDAPHLARHGVVVRGSMYASLTSNIEVYHGRYERDASRGYGLSAGSRVPF 1012
Db 960 TTDAPHLARHGVVVRGSMYASLTSNIEVYHGRYERDASRGYGLSAGSRVPF 1012

RESULT 21

US-10-931-779-15
; Sequence 15, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BPL04
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-15

Query Match 97.0%; Score 5106.5; DB 5; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 17; Mismatches 15; Indels 1; Gaps 1;

Qy 1 MQTSPHKFPLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF 60
Db 1 MQTSPHKFPLSMILAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF 60

Qy 61 SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITPENIRTSNGAALSADGL 120
Db 61 SAGELTLKLNLDNSIAALPLSCFNLGSGFTVLGRGHSITPENIRTSNGAALSADGL 120

Qy 121 FTIEGFKELSPNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180
Db 121 FTIEGFKELSPNCNLSLAVLPAATTNKGSTPTTTPSNGTIYSKTDLLLNNEKPSF 180

Qy 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTAQDGGACQVTFSPSAMANEAPAFVAN 240
Db 181 YSNLVSGDGGIDAISLTVQGISKLCVFOENTAQDGGACQVTFSPSAMANEAPAFVAN 240

Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAPL 300
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAPL 300

Qy 301 NNGKTLFLNNVASPVYIAAKQPTSOASNTSNYGDGGAIFCKNGAQ-AGSNNGSVSFD 359
Db 301 NNGKTLFLNNVASPVYIAAKQPTSOASNTSNYGDGGAIFCKNGAQ-AGSNNGSVSFD 359

Qy 360 GEGVVFSSNVAAGKGAIIYAKKLSVANCGPVQFLGNANDGGAIIYLGESGELSADYG 419
Db 360 GEGVVFSSNVAAGKGAIIYAKKLSVANCGPVQFLGNANDGGAIIYLGESGELSADYG 419

Qy 420 DIIFDGNLKRKTAKENAADVNGTVSSQAIISWGSGGKIITTLAKAGHQLFPNDPIEMANGN 479
Db 420 DIIFDGNLKRKTAKENAADVNGTVSSQAIISWGSGGKIITTLAKAGHQLFPNDPIEMANGN 479

Qy 480 NQPAQSSKLLKINDGEGVTGDIIVFANGSSTLYQNTVIEQGRIVLRKAKLSVNSISQGG 539
Db 480 NQPAQSSKLLKINDGEGVTGDIIVFANGSSTLYQNTVIEQGRIVLRKAKLSVNSISQGG 539

Qy 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLNLHLSSLANNANVTNPPNPPAQDSHP 599
Db 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLNLHLSSLANNANVTNPPNPPAQDSHP 599

Qy 600 AVIGSTTAGSVTIISGPIFFEDLDDTAYDRYDNLGSKINVLKQLGTPPANAPSDTL 659
Db 600 AVIGSTTAGSVTIISGPIFFEDLDDTAYDRYDNLGSKINVLKQLGTPPANAPSDTL 659

Qy 660 GNEAPKYGQSGWKLWDPNTANNPPTLKATWTGTGYNPCGPVRSIVLPSNLGSIIDI 719
Db 660 GNEAPKYGQSGWKLWDPNTANNPPTLKATWTGTGYNPCGPVRSIVLPSNLGSIIDI 719

Qy 720 RSAHSIAIQASVDGRSICRGLWVSGVSNFFYHDDRALGQGYRIISGGYSLGANSYFGSSMF 779
Db 720 RSAHSIAIQASVDGRSICRGLWVSGVSNFFYHDDRALGQGYRIISGGYSLGANSYFGSSMF 779

Qy 780 GLATEVFGRSKDYVVCRSNHACIGSVYLSQOALCGSYLFGDAFIRASVFGNGHMKT 839
Db 780 GLATEVFGRSKDYVVCRSNHACIGSVYLSQOALCGSYLFGDAFIRASVFGNGHMKT 839

Qy 840 SYTFAESDVRWDDNNCLAGIEGAGLPIVITPSKLYLNELRPFVQAEPSYADHESPTSEGD 899
Db 840 SYTFAESDVRWDDNNCLAGIEGAGLPIVITPSKLYLNELRPFVQAEPSYADHESPTSEGD 899

Qy 900 QARAPKSGHLLNLSPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTTTLLSHQBTW 959
Db 900 QARAPKSGHLLNLSPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTTTLLSHQBTW 959

Qy 960 TTDAPHLARHGVVVRGSMYASLTSNIEVYHGRYERDASRGYGLSAGSRVPF 1012
Db 960 TTDAPHLARHGVVVRGSMYASLTSNIEVYHGRYERDASRGYGLSAGSRVPF 1012

QY 480 NPAQSSKLLKINDGEGTGDIVFANGSTLYQNTTIEQGRIVLREKAKLSVNSLSQTGG 539
 DB 481 NPAQSSBPLKINDGEGTGDIVFANGSTLYQNTTIEQGRIVLREKAKLSVNSLSQTGG 540
 QY 540 SLYMEAGSTWDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNTPPPAQDSHP 599
 DB 541 SLYMEAGSTLDPVTPQPPQPPAANQSLITLSNLHLSLSLLANNAVTPNTPPPAQDSHP 600
 QY 600 AVIGSTTAGSVTISGPIPEDDDTAYDRYDGLWGSNQNKLVLQGLQTPPANAPSDTL 659
 DB 601 AVIGSTTAGSVTISGPIPEDDDTAYDRYDGLWGSNQNKLVLQGLQTPPANAPSDTL 660
 QY 660 GNEMPKYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVASFVNSLWGSILDI 719
 DB 661 GNEMPKYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVASFVNSLWGSILDI 720
 QY 720 RSHASAIQASVDRSVCRLWGSVSNFYHDDALGQGRYISGYSIGANSYFGSSMF 779
 DB 721 RSHASAIQASVDRSVCRLWGSVSNFYHDDALGQGRYISGYSIGANSYFGSSMF 780
 QY 780 GLAPTEVFGRSKDYVVCRRNHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKT 839
 DB 781 GLAPTEVFGRSKDYVVCRRNHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKT 840
 QY 840 SYTFAESDVRDNNCLAGEIGAGLPVITPSPKLYNELRPPVQAFPSYADHESFTEGD 899
 DB 841 SYTFAESDVRDNNCLAGEIGAGLPVITPSPKLYNELRPPVQAFPSYADHESFTEGD 900
 QY 900 QARAFKSHLNLVSPGVKFCRSTHNPKNYSFMAAYICDAVRTISGTETLLSHOETW 959
 DB 901 QARAFKSHLNLVSPGVKFCRSTHNPKNYSFMAAYICDAVRTISGTETLLSHOETW 960
 QY 960 TTDAPFLARHGVVVRGSMYASLTNSIEVYGHGREYERDASRGYLSAGSKVRFP 1012
 DB 961 TTDAPFLARHGVVVRGSMYASLTNSIEVYGHGREYERDASRGYLSAGSKVRFP 1013
 RESULT 22
 US-10-931-779-16
 ; Sequence 16, Application US/10931779
 ; Publication No. US20050048557A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jackson, W. James
 ; APPLICANT: Face, John
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
 ; FILE REFERENCE: RPI04
 ; CURRENT APPLICATION NUMBER: US/10/931,779
 ; CURRENT FILING DATE: 2004-09-01
 ; PRIOR APPLICATION NUMBER: 09/542,520
 ; PRIOR FILING DATE: 200-04-03
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 16
 ; LENGTH: 1013
 ; TYPE: PRF
 ; ORGANISM: Chlamydia sp.
 US-10-931-779-16
 Query Match 97.0%; Score 5106.5; DB 5; Length 1013;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;
 QY 1 MQTSFHKPFLSMILAYSCCLNGGVAEIMVPGQIYDGETLVSPFTYVIGDPSGTTVP 60
 DB 1 MQTSFHKPFLSMILAYSCCLNGGVAEIMVPGQIYDGETLVSPFTYVIGDPSGTTVP 60
 QY 61 SAGELTTLKNDNSIAALPLSCFNGLLGSFTVLGRGHSLTPEINTSTNGAALSNSAADGL 120
 DB 61 SAGELTTLKNDNSIAALPLSCFNGLLGSFTVLGRGHSLTPEINTSTNGAALSNSAADGL 120
 QY 121 FTIEGFKELSPNSCNLSLAVLPAATTNNGSQTPPTTTSTPSNGTIYSKTDLLLLANKEFSF 180

DB 121 FTIEGFKELSPNSCNLSLAVLPAATTNNGSQTPPTTTSTPSNGTIYSKTDLLLLANKEFSF 180
 QY 181 YSNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAOADGACQVWTSFSAMANEAPLAFVAN 240
 DB 181 YSNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAOADGACQVWTSFSAMANEAPLAFVAN 240
 QY 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVGSFRNTAYEPDGNVARVGGGIYSYGNVAF 300
 DB 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVGSFRNTAYEPDGNVARVGGGIYSYGNVAF 300
 QY 301 NNGKTLFLNNVASPVYIAAKOPTSOASNTSNYDGGAIKCKNGAQ-AGSNNSVSVD 359
 DB 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAIKCKNGAQAGSNNSVSVD 360
 QY 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFURNIANDGGAIYLGESGELSADYG 419
 DB 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFURNIANDGGAIYLGESGELSADYG 420
 QY 420 DIIFDGNLKRRTAKENAADVNGVTVSSQASISMGSGGKITTLLRAKAGHQLFNDPIEMANG 479
 DB 421 DIIFDGNLKRRTAKENAADVNGVTVSSQASISMGSGGKITTLLRAKAGHQLFNDPIEMANG 480
 QY 480 NPAQSSKLLKINDGEGTGDIVFANGSTLYQNTTIEQGRIVLREKAKLSVNSLSQTGG 539
 DB 481 NPAQSSBPLKINDGEGTGDIVFANGSTLYQNTTIEQGRIVLREKAKLSVNSLSQTGG 540
 QY 540 SLYMEAGSTWDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNTPPPAQDSHP 599
 DB 541 SLYMEAGSTLDPVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPNTPPPAQDSHP 600
 QY 600 AVIGSTTAGSVTISGPIPEDDDTAYDRYDGLWGSNQNKLVLQGLQTPPANAPSDTL 659
 DB 601 AVIGSTTAGSVTISGPIPEDDDTAYDRYDGLWGSNQNKLVLQGLQTPPANAPSDTL 660
 QY 660 GNEMPKYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVASFVNSLWGSILDI 719
 DB 661 GNEMPKYQGSWKLAWDPNTANNPGYTLKATWTKTGNPGRVASFVNSLWGSILDI 720
 QY 720 RSHASAIQASVDRSVCRLWGSVSNFYHDDALGQGRYISGYSIGANSYFGSSMF 779
 DB 721 RSHASAIQASVDRSVCRLWGSVSNFYHDDALGQGRYISGYSIGANSYFGSSMF 780
 QY 780 GLAPTEVFGRSKDYVVCRRNHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKT 839
 DB 781 GLAPTEVFGRSKDYVVCRRNHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHMKT 840
 QY 840 SYTFAESDVRDNNCLAGEIGAGLPVITPSPKLYNELRPPVQAFPSYADHESFTEGD 899
 DB 841 SYTFAESDVRDNNCLAGEIGAGLPVITPSPKLYNELRPPVQAFPSYADHESFTEGD 900
 QY 900 QARAFKSHLNLVSPGVKFCRSTHNPKNYSFMAAYICDAVRTISGTETLLSHOETW 959
 DB 901 QARAFKSHLNLVSPGVKFCRSTHNPKNYSFMAAYICDAVRTISGTETLLSHOETW 960
 QY 960 TTDAPFLARHGVVVRGSMYASLTNSIEVYGHGREYERDASRGYLSAGSKVRFP 1012
 DB 961 TTDAPFLARHGVVVRGSMYASLTNSIEVYGHGREYERDASRGYLSAGSKVRFP 1013

RESULT 23
 US-09-841-132-190
 ; Sequence 190, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probat, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599

Db	385	VQFLNIANDGGAIYLGSSGELSLADYDGIIFDGNLAKRTAKENAADVNGVTVSSQAISM	444	Qy	31	MVPGIYDGETLVTSFPYTVIGDPSGTTVPSAGBELTLKNLDNSIAALPLSCFGLNLSGFT	90
Qy	451	GSGGKITTLRAKAGHOIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL	510	Db	1	MIPOGIYDGETLVTSFPYTVIGDPSGTTVPSAGBELTLKNLDNSIAALPLSCFGLNLSGFT	60
Db	445	GSGGKITTLRAKAGHOIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL	504	Qy	91	VLGRHSLTFENIETSTNGAALSNSAADGLFTIIEGFKELSFNSCNLSLAVLPAATNKG	150
Qy	511	YQNTIEOGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPOPPQPPAANQILITLS	570	Db	61	VLGRHSLTFENIETSTNGAALSNSAADGLFTIIEGFKELSFNSCNLSLAVLPAATNKG	120
Db	505	YQNTIEOGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPOPPQPPAANQILITLS	564	Qy	151	QTPFTTSTPSNGTIYSKTDLLLLNNEKPSFYSNLVSGDGAIDAKSLTVQGISKLCPQOE	210
Qy	571	NLHLSLSLLANNAVTPNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD	630	Db	121	QTPFTTSTPSNGTIYSKTDLLLLNNEKPSFYSNLVSGDGAIDAKSLTVQGISKLCPQOE	180
Db	565	NLHLSLSLLANNAVTPNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD	624	Qy	211	NTAQADGACQVWTSFSAANEAPIAFVANVAGVRGGIAAVQDQOQGVSSSTSTEDPVV	270
Qy	631	WLGSNOKINVLKQLGKTPPANAPSDLTILGNEMPKYQGSWKLAWDPNTANNPPTLKA	690	Db	181	NTAQADGACQVWTSFSAANEAPIAFVANVAGVRGGIAAVQDQOQGVSSSTSTEDPVV	240
Db	625	WLGSNOKINVLKQLGKTPPANAPSDLTILGNEMPKYQGSWKLAWDPNTANNPPTLKA	684	Qy	271	SFSRNTAVFPGNVARVGGIYSYCNVAFNNGKTLFLNNVASPVYIAAKQPTSOASNT	330
Qy	691	TWTKTYNPGPERRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	750	Db	241	SFSRNTAVFPGNVARVGGIYSYCNVAFNNGKTLFLNNVASPVYIAAKQPTSOASNT	300
Db	685	TWTKTYNPGPERRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	744	Qy	331	SNNYGDGGAIFCKNGAQAGSNNSGSVSPDGEVGFSSNVAAGKGAIAKLSVANC	390
Qy	751	DRDALGQGYRYISGGYSLGANSYFGSSMFLAFTVEVFGRSKDYVVCRSNHHACIGSVYLS	810	Db	301	SNNYGDGGAIFCKNGAQAGSNNSGSVSPDGEVGFSSNVAAGKGAIAKLSVANC	360
Db	745	DRDALGQGYRYISGGYSLGANSYFGSSMFLAFTVEVFGRSKDYVVCRSNHHACIGSVYLS	804	Qy	391	VQFLNIANDGGAIYLGSSGELSLADYDGIIFDGNLAKRTAKENAADVNGVTVSSQAISM	450
Qy	811	TQOALCGSYLFGDAFIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGIAGLPIVITP	870	Db	361	VQFLNIANDGGAIYLGSSGELSLADYDGIIFDGNLAKRTAKENAADVNGVTVSSQAISM	420
Db	805	TQOALCGSYLFGDAFIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGIAGLPIVITP	864	Qy	451	GSGGKITTLRAKAGHOIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL	510
Qy	871	SKLYNELAPPVQAEBSYADHESFTEEGDQARAFKSGHLLNLSVPVGVKFDRCSSSTHPNK	930	Db	421	GSGGKITTLRAKAGHOIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL	480
Db	865	SKLYNELAPPVQAEBSYADHESFTEEGDQARAFKSGHLLNLSVPVGVKFDRCSSSTHPNK	924	Qy	511	YQNTIEOGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPOPPQPPAANQILITLS	570
Qy	931	YSFMAAICDAVRTISGTETTLTLLSHOETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	990	Db	481	YQNTIEOGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPOPPQPPAANQILITLS	540
Db	925	YSFMAAICDAVRTISGTETTLTLLSHOETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	984	Qy	571	NLHLSLSLLANNAVTPNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD	630
Qy	991	GRYEYRDASRGYGLSAGSVRFP 1012		Db	541	NLHLSLSLLANNAVTPNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYD	600
Db	985	GRYEYRDASRGYGLSAGSVRFP 1006		Qy	631	WLGSNOKINVLKQLGKTPPANAPSDLTILGNEMPKYQGSWKLAWDPNTANNPPTLKA	690
				Db	601	WLGSNOKINVLKQLGKTPPANAPSDLTILGNEMPKYQGSWKLAWDPNTANNPPTLKA	660
				Qy	691	TWTKTYNPGPERRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	750
				Db	661	TWTKTYNPGPERRVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	720
				Qy	751	DRDALGQGYRYISGGYSLGANSYFGSSMFLAFTVEVFGRSKDYVVCRSNHHACIGSVYLS	810
				Db	721	DRDALGQGYRYISGGYSLGANSYFGSSMFLAFTVEVFGRSKDYVVCRSNHHACIGSVYLS	780
				Qy	811	TQOALCGSYLFGDAFIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGIAGLPIVITP	870
				Db	781	TQOALCGSYLFGDAFIRASYGFGNQHMKTSTYTPAESDVRWNNCLAGIAGLPIVITP	840
				Qy	871	SKLYNELAPPVQAEBSYADHESFTEEGDQARAFKSGHLLNLSVPVGVKFDRCSSSTHPNK	930
				Db	841	SKLYNELAPPVQAEBSYADHESFTEEGDQARAFKSGHLLNLSVPVGVKFDRCSSSTHPNK	900
				Qy	931	YSFMAAICDAVRTISGTETTLTLLSHOETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	990
				Db	901	YSFMAAICDAVRTISGTETTLTLLSHOETWTTDAFLARHGTVVVRGSMYASLTSNIEVYGH	960
				Qy	991	GRYEYRDASRGYGLSAGSVRFP 1012	
				Db	961	GRYEYRDASRGYGLSAGSVRFP 982	

RESULT 26
US-10-872-155-176
; Sequence 176, Application US/10872155
; Publication No. US20040234536A1

Query Match 96.5%; Score 5084; DB 3; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 25
US-09-841-132-176
; Sequence 176, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-841-132-176

```

; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C10
; CURRENT APPLICATION NUMBER: US/10/872,155
; CURRENT FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 09/620,412
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 09/598,419
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/556,877
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/454,684
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 09/426,571
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 09/410,568
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/288,594
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 09/208,277
; PRIOR FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-872-155-176

Query Match      96.5%; Score 5084; DB 5; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 31 MVPGIYDGETLTVSFPYTVIGDPSGTTVFAGELTLKNLNSIAALPLSCFNGLLGSFT 90
Db 1 MIPQIYDGETLTVSFPYTVIGDPSGTTVFAGELTLKNLNSIAALPLSCFNGLLGSFT 60

Qy 91 VLGRGHSLTPEIRITSTNGAALSNSAADGLFTIEGKELSFNSCNLSLAVLPAATTNKS 150
Db 61 VLGRGHSLTPEIRITSTNGAALSNSAADGLFTIEGKELSFNSCNLSLAVLPAATTNKS 120

Qy 151 QTPTTSTPSNGTIYSKTDLLILNNEKESFYSLNLSVSGDGAIDAKSLTVQGISKLCVFOE 210
Db 121 QTPTTSTPSNGTIYSKTDLLILNNEKESFYSLNLSVSGDGAIDAKSLTVQGISKLCVFOE 180

Qy 211 NTAQADGACQVWTSFSAMANEAPLAFVANAVRGGGIAAVQDQOQGVSSSTSTEDPVV 270
Db 181 NTAQADGACQVWTSFSAMANEAPLAFVANAVRGGGIAAVQDQOQGVSSSTSTEDPVV 240

Qy 271 SFSRNTAVEPDGNVARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGQASNT 330
Db 241 SFSRNTAVEPDGNVARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGQASNT 300

Qy 331 SNNTYDGGAIKCKGQAQAGNSNSGVSFDGEGVVPFSSNVAAGKGAIAKSLVANGCP 390
Db 301 SNNTYDGGAIKCKGQAQAGNSNSGVSFDGEGVVPFSSNVAAGKGAIAKSLVANGCP 360

Qy 391 VOFLRNANDGAIYLGESGELSADYGDIIFDGNLAKRTAKENAADVNGTVVSSQALSM 450
Db 361 VOFLRNANDGAIYLGESGELSADYGDIIFDGNLAKRTAKENAADVNGTVVSSQALSM 420

Qy 451 GSGGKITTLRACAGHQLIFNDPIEMANGNQPQSSKLLKINDGEGYTGDIVFANGSSTL 510
Db 421 GSGGKITTLRACAGHQLIFNDPIEMANGNQPQSSKLLKINDGEGYTGDIVFANGSSTL 480

; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C3
; CURRENT APPLICATION NUMBER: US/10/197,220
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 169
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Chlamydia
US-10-197-220-169

Query Match      63.3%; Score 3336.5; DB 5; Length 670;
Best Local Similarity 97.2%; Pred. No. 3.6e-238;
Matches 651; Conservative 11; Mismatches 7; Indels 1; Gaps 1;

Qy 28 AEIMVPGIYDGETLTVSFPYTVIGDPSGTTVFAGELTLKNLNSIAALPLSCFNGLLG 87
Db 1 AEIMVPGIYDGETLTVSFPYTVIGDPSGTTVFAGELTLKNLNSIAALPLSCFNGLLG 60

Qy 88 SPTVLGRGHSLTPEIRITSTNGAALSNSAADGLFTIEGKELSFNSCNLSLAVLPAATTN 147
Db 61 SPTVLGRGHSLTPEIRITSTNGAALSNSAADGLFTIEGKELSFNSCNLSLAVLPAATTN 120

Qy 148 KGSQTPTTSTPSNGTIYSKTDLLILNNEKESFYSLNLSVSGDGAIDAKSLTVQGISKLCV 207
Db 121 NGSQTPTTSTPSNGTIYSKTDLLILNNEKESFYSLNLSVSGDGAIDAKSLTVQGISKLCV 180
```

208 POENTAOADGGACQCVVTSFSAMANEAPAFVANVAGVGGGIAAVODGCGVSSSTSTED 267
181 FOENTAOADGGACQCVVTSFSAMANEAPAFVANVAGVGGGIAAVODGCGVSSSTSTED 240
268 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAKQPTSGQA 327
241 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAEQPTNGQA 300
328 SNTSNNGYDGGAI FCKNGAQ-AGSNNSGVSFDDGCVVFFSSNVAAGKGGAIYAKKLSVA 386
301 SNTSDNYDGGAI FCKNGAQAGSNNSGVSFDDGCVVFFSSNVAAGKGGAIYAKKLSVA 360
387 NCGPVQFLNNTANDGGAIVLGGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 446
361 NCGPVQFLNNTANDGGAIVLGGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 420
447 AISMGGGKITTILRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFANG 506
421 AISMGGGKITTILRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFANG 480
507 SSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPPOPPQPPAANQL 566
481 NSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPPOPPQPPAANQL 540
567 ITLSNLHLSLSLLANNVNTPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626
541 ITLSNLHLSLSLLANNVNTPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAY 600
627 DRYDWMGSKNQKIDVLKQLGTQPSANAPSDLTGLNEMPKYGYGGSWKLAWDPNTANNPY 686
601 DRYDWMGSKNQKIDVLKQLGTQPSANAPSDLTGLNEMPKYGYGGSWKLAWDPNTANNPY 660
687 TLKATWTKTG 696
661 TLKATWTKTG 670

RESULT 28
US-11-109-468-169
; Sequence 169, Application US/11109468
; Publication No. US20050232941A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Gunderian, Jeff
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Maisonneuve, Jean-Francois L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C4
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US 10/197,220
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 10/007,693
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 10/012,256
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 09/841,260
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/219,752
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/198,853
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 169
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Chlamydia
US-11-109-468-169

Query Match 63.3%; Score 3336.5; DB 6; Length 670;
Best Local Similarity 97.2%; Pred. No. 3.6e-238;

Matches 651; Conservative 11; Mismatches 7; Indels 1; Gaps 1;
QY 28 AEIMVQGIYDGETLTVSFPVTVIGDPSGTTVFAGELTTLKNLNSIAALPLSCFNNLLG 87
DB 1 AEIMVQGIYDGETLTVSFPVTVIGDPSGTTVFAGELTTLKNLNSIAALPLSCFNNLLG 60
QY 88 SFTVLGRGHSLTFFENIRTSNNGAALSADGLFTIEGFKELSPNSCNLSLAVLPAATN 147
DB 61 SFTVLGRGHSLTFFENIRTSNNGAALSADGLFTIEGFKELSPNSCNLSLAVLPAATN 120
QY 148 KGSQPTTSTPSNGTIYSKTDLLILNNEKESFYNLSVGGGADAKSLTVQGISKLCV 207
DB 121 NGSQPTTSTPSNGTIYSKTDLLILNNEKESFYNLSVGGGADAKSLTVQGISKLCV 180
QY 208 FOENTAOADGGACQCVVTSFSAMANEAPAFVANVAGVGGGIAAVODGCGVSSSTSTED 267
DB 181 FOENTAOADGGACQCVVTSFSAMANEAPAFVANVAGVGGGIAAVODGCGVSSSTSTED 240
QY 268 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAKQPTSGQA 327
DB 241 PVVSFSRNTAVFEDGNVAVGGGIYSYGNVAFNNNGKTLFLNNVASPVYIAAEQPTNGQA 300
QY 328 SNTSNNGYDGGAI FCKNGAQ-AGSNNSGVSFDDGCVVFFSSNVAAGKGGAIYAKKLSVA 386
DB 301 SNTSDNYDGGAI FCKNGAQAGSNNSGVSFDDGCVVFFSSNVAAGKGGAIYAKKLSVA 360
QY 387 NCGPVQFLNNTANDGGAIVLGGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 446
DB 361 NCGPVQFLNNTANDGGAIVLGGESGELSADYGDII FDNLKRKTAKENAAADVNGVTSSQ 420
QY 447 AISMGGGKITTILRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFANG 506
DB 421 AISMGGGKITTILRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFANG 480
QY 507 SSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPPOPPQPPAANQL 566
DB 481 NSTLYQNVTEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPPOPPQPPAANQL 540
QY 567 ITLSNLHLSLSLLANNVNTPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626
DB 541 ITLSNLHLSLSLLANNVNTPTNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAY 600
QY 627 DRYDWMGSKNQKIDVLKQLGTQPSANAPSDLTGLNEMPKYGYGGSWKLAWDPNTANNPY 686
DB 601 DRYDWMGSKNQKIDVLKQLGTQPSANAPSDLTGLNEMPKYGYGGSWKLAWDPNTANNPY 660
QY 687 TLKATWTKTG 696
DB 661 TLKATWTKTG 670

RESULT 29
US-10-701-844-17
; Sequence 17, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-701-844-17


```
Query Match      48.5%; Score 2552; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.3e-180;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGIYDGETLTVSPFYTIVGDPGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 88
DB 1 EIMVPGIYDGETLTVSPFYTIVGDPGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 60

QY 89 FTVLGRGHSITPENIRTSNKAALSNSAADGLFTIEGFKELSPFNCNLSLLAVLPAATTNK 148
DB 61 FTVLGRGHSITPENIRTSNKAALSNSAADGLFTIEGFKELSPFNCNLSLLAVLPAATTNK 120

QY 149 GSQPTTSTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVP 208
DB 121 GSQPTTSTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVP 180

QY 269 VVFSRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328
DB 241 VVFSRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

QY 329 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 388
DB 301 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 360

QY 389 GPVQFLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 448
DB 361 GPVQFLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

QY 449 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 508
DB 421 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
DB 481 TLYQNVTIEQGRIVLREKAKLSVNS 505

RESULT 30
US-10-766-711-17
; Sequence 17, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-766-711-17

Query Match      48.5%; Score 2552; DB 4; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.3e-180;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGIYDGETLTVSPFYTIVGDPGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 88
DB 1 EIMVPGIYDGETLTVSPFYTIVGDPGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 60

QY 89 FTVLGRGHSITPENIRTSNKAALSNSAADGLFTIEGFKELSPFNCNLSLLAVLPAATTNK 148
DB 61 FTVLGRGHSITPENIRTSNKAALSNSAADGLFTIEGFKELSPFNCNLSLLAVLPAATTNK 120

QY 149 GSQPTTSTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVP 208
DB 121 GSQPTTSTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVP 180

QY 269 VVFSRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328
DB 241 VVFSRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

QY 329 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 388
DB 301 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 360

QY 389 GPVQFLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 448
DB 361 GPVQFLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

QY 449 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 508
DB 421 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
DB 481 TLYQNVTIEQGRIVLREKAKLSVNS 505

RESULT 31
US-10-931-779-17
; Sequence 17, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: BPI04
; CURRENT APPLICATION NUMBER: US/10/931,779
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 09/542,520
; PRIOR FILING DATE: 200-04-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-10-931-779-17

Query Match      48.5%; Score 2552; DB 5; Length 505;
Best Local Similarity 100.0%; Pred. No. 3.3e-180;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 EIMVPGIYDGETLTVSPFYTIVGDPGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 88
DB 1 EIMVPGIYDGETLTVSPFYTIVGDPGTTVFSAGELTLKLNLSIAALPLSCFNLGSS 60

QY 89 FTVLGRGHSITPENIRTSNKAALSNSAADGLFTIEGFKELSPFNCNLSLLAVLPAATTNK 148
DB 61 FTVLGRGHSITPENIRTSNKAALSNSAADGLFTIEGFKELSPFNCNLSLLAVLPAATTNK 120

QY 149 GSQPTTSTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVP 208
DB 121 GSQPTTSTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVP 180

QY 209 QENTAQADGGACQVVTTSFSAANEAPAFVANVAVGSGGIAAVQDQGGQVSSSTSTEDP 268
DB 181 QENTAQADGGACQVVTTSFSAANEAPAFVANVAVGSGGIAAVQDQGGQVSSSTSTEDP 240

QY 269 VVFSRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 328
DB 241 VVFSRNTAVSFDCGNVARVGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300

QY 329 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 388
DB 301 NTSNNYGDGGAIFCKNGAQAGSNNSGVSFDGEGVVFSSNVAAGKGGAIYAKKLSVANC 360

QY 389 GPVQFLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 448
DB 361 GPVQFLRNANDGGAIIYLGESGELSADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAI 420

QY 449 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 508
DB 421 SMGSGGKITTLRAKAGHOLLFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVPANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
DB 481 TLYQNVTIEQGRIVLREKAKLSVNS 505
```

Db 361 SSSLANNVTPPTNPQAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 420
QY 637 KINVLKQLQGTGKPPANAPSDTLTGNEMPKYGYQGSWKL 674
Db 421 KINVLKQLQGTGKPPANAPSDTLTGNEMPKYGYQGSWKL 458

RESULT 33
US-10-766-711-36
; Sequence 36, Application US/10766711
; Publication No. US20040137005A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-062-999
; CURRENT APPLICATION NUMBER: US/10/766,711
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US/08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; TYPE: PRT
; LENGTH: 458
; ORGANISM: Chlamydia sp.
US-10-766-711-36

Query Match 44.6%; Score 2350; DB 4; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.7e-165;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GCACQVVTSPFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVSFSRNT 276
Db 1 GCACQVVTSPFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVSFSRNT 60

QY 277 AVEFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 336
Db 61 AVEFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 120

QY 337 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANCGPVQFLRN 396
Db 121 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANCGPVQFLRN 180

QY 397 IANDGGAIIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISMGSGGKI 456
Db 181 IANDGGAIIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISMGSGGKI 240

QY 457 TTLRAKAGHQIILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 516
Db 241 TTLRAKAGHQIILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 300

QY 517 EQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 576
Db 301 EQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 360

QY 577 SSSLANNVTPPTNPQAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 636
Db 361 SSSLANNVTPPTNPQAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 420

QY 637 KINVLKQLQGTGKPPANAPSDTLTGNEMPKYGYQGSWKL 674
Db 421 KINVLKQLQGTGKPPANAPSDTLTGNEMPKYGYQGSWKL 458

RESULT 34
US-10-931-779-36
; Sequence 36, Application US/10931779
; Publication No. US20050048557A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof

Db 241 VVSFERTAVFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS 300
QY 329 NTSNNYGDGGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANC 388
Db 301 NTSNNYGDGGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANC 360

QY 389 GPVQFLRNTANDGGAIIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAI 448
Db 361 GPVQFLRNTANDGGAIIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAI 420

QY 449 SMGSGGKITTLRAKAGHQIILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSS 508
Db 421 SMGSGGKITTLRAKAGHQIILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSS 480

QY 509 TLYQNVTIEQGRIVLREKAKLSVNS 533
Db 481 TLYQNVTIEQGRIVLREKAKLSVNS 505

RESULT 32
US-10-701-844-36
; Sequence 36, Application US/10701844
; Publication No. US20040067524A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/10/701,844
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/612,402B
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; TYPE: PRT
; LENGTH: 458
; ORGANISM: Chlamydia sp.
US-10-701-844-36

Query Match 44.6%; Score 2350; DB 4; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.7e-165;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GCACQVVTSPFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVSFSRNT 276
Db 1 GCACQVVTSPFSAMANEAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVSFSRNT 60

QY 277 AVEFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 336
Db 61 AVEFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGD 120

QY 337 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANCGPVQFLRN 396
Db 121 GGAIFCKNGAQAQSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANCGPVQFLRN 180

QY 397 IANDGGAIIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISMGSGGKI 456
Db 181 IANDGGAIIYLGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISMGSGGKI 240

QY 457 TTLRAKAGHQIILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 516
Db 241 TTLRAKAGHQIILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 300

QY 517 EQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 576
Db 301 EQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 360

QY 577 SSSLANNVTPPTNPQAQDHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLGSNQ 636

```
/ FILE REFERENCE: BP104
/ CURRENT APPLICATION NUMBER: US/10/931,779
/ CURRENT FILING DATE: 2004-09-01
/ PRIOR APPLICATION NUMBER: 09/542,520
/ PRIOR FILING DATE: 200-04-03
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 36
/ LENGTH: 458
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-10-931-779-36

Query Match      44.6%; Score 2350; DB 5; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.7e-185;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GCACQVVTSPFSAMANEAFIAFVANVAGVGGGIAAVQDQGVSSSTSTEDPVVPSRNT 276
DB 1 GCACQVVTSPFSAMANEAFIAFVANVAGVGGGIAAVQDQGVSSSTSTEDPVVPSRNT 60

QY 277 AVFPGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 336
DB 61 AVFPGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNNGD 120

QY 337 GGAIFCKGAQAGSNNSGSVPDGEVVPFSSNVAAGGAIYAKKLSVANCGPVQFLRN 396
DB 121 GGAIFCKGAQAGSNNSGSVPDGEVVPFSSNVAAGGAIYAKKLSVANCGPVQFLRN 180

QY 397 IANDGAIYLGESGELSLADYDGIIFDGNLKRTAKENAADVNGVTVSQAISMGGGKI 456
DB 181 IANDGAIYLGESGELSLADYDGIIFDGNLKRTAKENAADVNGVTVSQAISMGGGKI 240

QY 457 TTLAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 516
DB 241 TTLAKAGHQLFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTLYQNVTI 300

QY 517 EGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPQPPAAANQLITLSNHL 576
DB 301 EGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPQPPAAANQLITLSNHL 360

QY 577 SLLANNAVTPNTPPAQDHPAVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQ 636
DB 361 SLLANNAVTPNTPPAQDHPAVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQ 420

QY 637 KINVLKQLGKTPKANAPSDLTGLNEMPKYGYQGSWKL 674
DB 421 KINVLKQLGKTPKANAPSDLTGLNEMPKYGYQGSWKL 458
```

```
RESULT 35
US-10-701-844-37
/ Sequence 37, Application US/10701844
/ Publication No. US20040067524A1
/ GENERAL INFORMATION:
/ APPLICANT: Jackson, W. James
/ TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
/ FILE REFERENCE: 7969-086-999
/ CURRENT APPLICATION NUMBER: US/10/701,844
/ CURRENT FILING DATE: 2003-11-04
/ PRIOR APPLICATION NUMBER: US/09/612,402B
/ PRIOR FILING DATE: 2000-07-06
/ PRIOR APPLICATION NUMBER: 08/942,596
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 37
/ LENGTH: 325
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-10-701-844-37
```

```
Query Match      32.9%; Score 1735; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.2e-120;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 747
DB 1 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 60

QY 748 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNHHACIGSV 807
DB 61 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNHHACIGSV 120

QY 808 YLSTQQAALCGSYLPGDAPIRASVYFGNQHMKTSYTFABESDVRWNNCLAGEICAGLP 927
DB 121 YLSTQQAALCGSYLPGDAPIRASVYFGNQHMKTSYTFABESDVRWNNCLAGEICAGLP 180

QY 868 ITPSKLYLNELRPPVQAEFSYADHESFTBEGDQARAFKSGHLLNLSVPVGVKFPDRCSSTH 927
DB 181 ITPSKLYLNELRPPVQAEFSYADHESFTBEGDQARAFKSGHLLNLSVPVGVKFPDRCSSTH 240

QY 928 PNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 987
DB 241 PNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 300

QY 988 YGHGRYEVDRASRGYGLSAGSRVRF 1012
DB 301 YGHGRYEVDRASRGYGLSAGSRVRF 325

RESULT 36
US-10-766-711-37
/ Sequence 37, Application US/10766711
/ Publication No. US20040137005A1
/ GENERAL INFORMATION:
/ APPLICANT: Jackson, W. James
/ APPLICANT: Pace, John
/ TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
/ FILE REFERENCE: 7969-062-999
/ CURRENT APPLICATION NUMBER: US/10/766,711
/ CURRENT FILING DATE: 2004-01-27
/ PRIOR APPLICATION NUMBER: US/08/942,596
/ PRIOR FILING DATE: 1997-10-02
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 37
/ LENGTH: 325
/ TYPE: PRT
/ ORGANISM: Chlamydia sp.
US-10-766-711-37
```

```
Query Match      32.9%; Score 1735; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.2e-120;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 688 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 747
DB 1 LKATWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNF 60

QY 748 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNHHACIGSV 807
DB 61 FYHDDALGQGYRYISGGYISLGANSYFGSSMFGLAFTFVFGRSKDYVVCRSNHHACIGSV 120

QY 808 YLSTQQAALCGSYLPGDAPIRASVYFGNQHMKTSYTFABESDVRWNNCLAGEICAGLP 927
DB 121 YLSTQQAALCGSYLPGDAPIRASVYFGNQHMKTSYTFABESDVRWNNCLAGEICAGLP 180

QY 868 ITPSKLYLNELRPPVQAEFSYADHESFTBEGDQARAFKSGHLLNLSVPVGVKFPDRCSSTH 927
DB 181 ITPSKLYLNELRPPVQAEFSYADHESFTBEGDQARAFKSGHLLNLSVPVGVKFPDRCSSTH 240

QY 928 PNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 987
DB 241 PNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHLARHGTVVVRGSMYASLTSNIEV 300
```


Query Match 21.4%; Score.1128; DB 3; Length 928;
Best Local Similarity 29.2%; Pred.No.2.8e-74;
Matches 311; Conservative 169; Mismatches 394; Indels 192; Gaps 32;

Qy 1 MOTSFHKPFLSMILAYSCCLNGGGYAAE-IMVPOQIYDGETLTWSPFYTVIGDPSGTVV 59
Db 1 MKSSPKFVFSTPAIFPLSMI-----ATETVLDSSASPDGNK-NGNFSVRESQEDAGTTY 54
Qy 60 FSAGELTLKVLNLSIAALPLSCFNLGSLFTVLGRGHSITPENIRTS-NGAALNSAAD 118
Db 55 LFKGNVTLNIPGTGTAITKSCFNNTKGLTFTGNGSNLLFQTVADGTVAGAAVNSVVVD 114
Qy 119 GLFTIEGPKELFSNCNSLLAVLPAATTNKGSTPTTTTPSNGTIYSKTDLLLLNNEKF 178
Db 115 KSTFTFGSLSP-----IASPGSSITTKGAVSCSTGS-----LSLTKNVL 157
Qy 179 SFYSNLVSGDGAIDAKSLTVQISKLCVQENTAQDGGACQVTVFSAMANEAPIAFV 238
Db 158 LFSKNFSTDNGGAIATKTLSTLTMSALFSENTSSKGGAIQTSDALITIGNQGEVSF- 216
Qy 239 ANVAGVRGGGIAAVDQOQGVSSSTSTEDPVVSFSRNTAVEPDGNVARVGGGIYSYGNVA 298
Db 217 -----SDNTSSDS-----GAAIFTEASVT 235
Qy 299 FLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNYYDGGAIKCKNGAQAGSNNSGVSP 358
Db 236 ISNNAKVSFDN-----KVTGASSSTTGD- SGGAI CAYK-----TSTDTKVTL 278
Qy 359 DGGVVPFSSNVAAGKGAIYAKLSVANGCPVQFLRNIA-----DGGAIYLGESGELS 413
Db 279 TGNQMLFNSNTTGTAGAIYVKLELASGELTLFSRNSVNGGTAPKGGAIATEDSGELS 338
Qy 414 LSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAISWGGCKITTLRAXAGHOILFNDPI 473
Db 339 LSADSGDIVPLGN---TVTSTTPCTN-----RGSIDLGTSAKMTALRSAAGRAIYFYDPI 390
Qy 474 EMANGNNQPAQSSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNTV 515
Db 391 TTGSSTT-----VTDVLKVNETPADSALQYTGNIIFTGEKLSATEAADSKNLTSLKLPVT 446
Qy 516 IEGRIVLREKAKLSVNSLSQTGS-LYMEAGSTWDFVTPPPQPPAANQLITLGNLHL 574
Db 447 LSGGTLSLKHGVTLQTAFTQQADSRLEMDVGITL-----PADTS--TINNLI 494
Qy 575 SLSSLLANNAVTPPTPPAQDGHAPVIGS--TTAGSVTISGPIFFEDLDLTAYDRYDWLG 633
Db 495 NISSI-----DGAKKAKIETKATSKNLTSLGTTITLLDPTGTFYENHS-LR 538
Qy 634 SNQKINVLKLOL-GTKPPANAPSLDTLGNEMPKYQGSW-KLAWDPNTANNGPYTLKAT 691
Db 539 NPQSYDILELKASGTVTSTAVTPDPIMGEKF-HYGYQGTWGPVIMGTGASTTATF-----N 593
Qy 692 WTKTYNPGPERVASLPNSLWGSILDIRSAHSAIAQASVDGRSYCRGLWVSGVSNPFYHD 751
Db 594 WTKTYIPNPERIGSLVPSLWNAFIDISLHYLMETANEGLQGDRAFCWAGLSNPFHKD 653
Qy 752 RDALGQGVRYISGGYSLGANSYFGS-SMFGLAFTVEFGRSKDYVVCRSNHHACIGSYLS 810
Db 654 STKTRGRFRLHSGYVIGGNLHCTCDKLSAFCQLFGDRDRDYFAKNQ-----GTVYGG 708
Qy 811 TQOALCGSYLFGDAFI-----RASYGFGNQHMKTSYTFABE 846
Db 709 TLX-----YQHNETVYISLPCKLRPCSLSYVPTETPVLPSGNLSYTHTDNDLKTYYTPT 763
Qy 847 SDVWDDNCLAGEIGAGLPVITPSKLYLNLRLPFOABESYADHESFTREGDOARAFKS 906
Db 764 VKSGWGNDSFALBFGGRAPICLDESALEP-EQYMPFMKLQFYAHQEGFKEQGEAREFGS 822
Qy 907 GHLNLVSVPGVKPDRCSSTHPNKYSFMAAYICDAYRTISGTETTLTSHQETTTTDAFHL 966
Db 823 SRLVNLALPIGIRFDKSDCQDATYNLTGLTVDLVRSNPDCTTTLRLISGDSWKTPTGNL 882

Search completed: May 13, 2006, 12:35:01
Job time : 176 secs

Qy 967 ARHGVVVRGSMYASLTNSNIEVYGHGRVEYRDASRGYGLSAGSRVRF 1012
Db 883 ARQALVLRAGNHFCFNSNFAPSQSPFELRGSSRNYNVDLGAKYQF 928

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:32:24 ; Search time 30 Seconds
(without alignments)
1583.741 Million cell updates/sec

Title: US-10-701-844-2

Perfect score: 5267

Sequence: 1 MQTSFHKKPLSMILAYSCS.....YERDASRGYLSAGSRVRF 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 250354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /SID55/prodata/2/pubpaa/US08_NEW_PUB.pep1.*
- 2: /SID55/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /SID55/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /SID55/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /SID55/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /SID55/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 7: /SID55/prodata/2/pubpaa/US09_NEW_PUB.pep1.*
- 8: /SID55/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 9: /SID55/prodata/2/pubpaa/US11_NEW_PUB.pep1.*
- 10: /SID55/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 11: /SID55/prodata/2/pubpaa/US11_NEW_PUB.pep1.*
- 12: /SID55/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5131.5	97.4	1013	11	US-11-103-957-9
2	5131.5	97.4	1013	11	US-11-018-868-21
3	771.5	14.6	878	11	US-11-103-957-7
4	771.5	14.6	878	11	US-11-018-868-23
5	649.5	12.3	964	11	US-11-103-957-13
6	649.5	12.3	964	11	US-11-018-868-19
7	615	11.7	1016	11	US-11-103-957-41
8	615	11.7	1016	11	US-11-018-868-22
9	597.5	11.3	1531	11	US-11-103-957-15
10	597.5	11.3	1531	11	US-11-018-868-18
11	551	10.5	975	11	US-11-103-957-45
12	551	10.5	975	11	US-11-018-868-15
13	551	10.5	975	11	US-11-018-868-50
14	522.5	9.9	1034	11	US-11-103-957-11
15	522.5	9.9	1034	11	US-11-018-868-20
16	475.5	9.0	1751	11	US-11-103-957-45
17	475.5	9.0	1751	11	US-11-018-868-16
18	412	7.8	1770	11	US-11-103-957-21
19	412	7.8	1770	11	US-11-018-868-17
20	237.5	4.5	1268	11	US-11-052-554A-1
21	232.5	4.4	1643	11	US-11-052-554A-172

22	228	4.3	1250	11	US-11-052-554A-16	Sequence 16, Appl
23	218	4.1	1571	11	US-11-052-554A-2	Sequence 2, Appl
24	218	4.1	3132	11	US-11-087-099-1245	Sequence 1245, Ap
25	218	4.1	3194	11	US-11-052-554A-90	Sequence 90, Appl
26	213	4.0	5291	11	US-11-052-554A-281	Sequence 281, App
27	204.5	3.9	2902	11	US-11-052-554A-91	Sequence 91, Appl
28	194	3.7	1474	11	US-11-067-260-18	Sequence 18, Appl
29	193.5	3.7	2340	11	US-11-052-554A-171	Sequence 171, App
30	193	3.7	1468	9	US-10-467-657-1088	Sequence 1088, Ap
31	193	3.7	2399	11	US-11-052-554A-92	Sequence 92, Appl
32	190	3.6	955	11	US-11-052-554A-179	Sequence 179, App
33	190	3.6	1392	11	US-11-067-260-24	Sequence 24, Appl
34	190	3.6	1447	11	US-11-067-260-22	Sequence 22, Appl
35	189.5	3.6	1461	11	US-11-052-554A-283	Sequence 283, App
36	189.5	3.6	1981	11	US-11-045-208-38	Sequence 38, Appl
37	189.5	3.6	2015	11	US-11-052-554A-374	Sequence 374, App
38	188.5	3.6	1312	11	US-11-067-260-20	Sequence 20, Appl
39	187.5	3.6	949	11	US-11-052-554A-6	Sequence 6, Appl
40	186.5	3.5	1343	11	US-11-052-554A-284	Sequence 284, App
41	184.5	3.5	1574	11	US-11-212-443-179	Sequence 179, App
42	184.5	3.5	1978	11	US-11-212-443-60	Sequence 60, Appl
43	182	3.5	596	8	US-10-196-749-310	Sequence 310, Appl
44	182	3.5	596	9	US-10-063-703-100	Sequence 100, App
45	182	3.5	596	9	US-10-194-487-310	Sequence 310, App

ALIGNMENTS

RESULT 1

US-11-103-957-9
; Sequence 9, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-9

Query Match 97.4%; Score 5131.5; DB 11; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY	1	MQTSFHKKPLSMILAYSCSLNGGGYAAIMVPOGLYDGETLVTFPPYTVIGDPGSGTTFV	60
Db	1	MQTSFHKKPLSMILAYSCSLNGGGYAAIMVPOGLYDGETLVTFPPYTVIGDPGSGTTFV	60
QY	61	SAGELTKNLNDNSIAALPLSCFNLGSGFTVLGRGHSITFFENIRTSNKAALNSAAGL	120
Db	61	SAGELTKNLNDNSIAALPLSCFNLGSGFTVLGRGHSITFFENIRTSNKAALSDSANSGL	120
QY	121	FTIEGKELSFNCNCSLLAVLPAATYTNKGSQPTTTSTPSNGTIYSKTDLLILNNEKFSF	180
Db	121	FTIEGKELSFNCNCSLLAVLPAATYTNKGSQPTTTSTPSNGTIYSKTDLLILNNEKFSF	180
QY	181	YSNLVSGDGAIDAKSLTVQGISKLVCVFQENTAQDGGACQVVTFSAMANEAPITAFVAN	240

```

181 YSNLVSGDGGAIADAKSLTVQGISKLCVFQENTAQDGGACQVTSFSAMANEAPAFIAN 240
241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
301 NGKTLFLNNVSPVYIAAKQPTSGQASNTNNYDGGGAIKFCXGQAG-AGNNNGSVSPD 359
301 NGKTLFLNNVSPVYIAAEQPTNGQASNTSDNYDGGGAIKFCXGQAG-AGNNNGSVSPD 360
360 GEGVVPFSSNAAGKGGAIYAKLSVANCGPVQFLNNTANDGGAIIYLGESGELSADYG 419
361 GEGVVPFSSNAAGKGGAIYAKLSVANCGPVQFLNNTANDGGAIIYLGESGELSADYG 420
420 DIIIDGNLKRKTAKENAADVNGTVSSQAI SMGSGGKITTLRAKAGHQLLPNDPIEMANGN 479
421 DIIIDGNLKRKTAKENAADVNGTVSSQAI SMGSGGKITTLRAKAGHQLLPNDPIEMANGN 480
480 NQPAQSSKLLKINDGEGTGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQGG 539
481 NQPAQSSKLLKINDGEGTGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQGG 540
540 SLYNEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 599
541 SLYNEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 600
600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 659
601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 660
660 GNEPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGYNPGRVVASI VPSNLSMGSILDI 719
661 GNEPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGYNPGRVVASI VPSNLSMGSILDI 720
720 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHRRDALGQGYRIYSGGYSLGANSYFGSSMF 779
721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHRRDALGQGYRIYSGGYSLGANSYFGSSMF 780
780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVYFGNQHMKT 839
781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVYFGNQHMKT 840
840 SYTFABESDVRWDNNCLVGEIGVGLPIVITPSKLYLNELRPFVQAEFSAHDESFTBEGD 899
841 SYTFABESDVRWDNNCLVGEIGVGLPIVITPSKLYLNELRPFVQAEFSAHDESFTBEGD 900
900 QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTQTLLSHQETW 959
901 QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTQTLLSHQETW 960
960 TTDAPFLARHGVIYRGSMYASLTSNIEYVGHGRVEYRDSRGYGLSAGSVRF 1012
961 TTDAPFLARHGVIYRGSMYASLTSNIEYVGHGRVEYRDSRGYGLSAGSVRF 1013
```

```

RESULT 2
US-11-018-868-21
; Sequence 21, Application US/1101868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guilio
; APPLICANT: Ratti, Guilio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
```

```

; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-21
Query Match 97.4%; Score 5131.5; DB 11; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;
Qy 1 MQTSFHFFFLSMILAYSCCSLNGGGYAAIRIMVPGIYDGETLTVSPFYTYTIGDPSGTTVF 60
Db 1 MQTSFHFFFLSMILAYSCCSLNGGGYAAIRIMVPGIYDGETLTVSPFYTYTIGDPSGTTVF 60
Qy 61 SAGELTLKLNLDNSIAALPLSCFNLGLSFTVLGRGHSHTFENIRTSNGAALSAAADGL 120
Db 61 SAGELTLKLNLDNSIAALPLSCFNLGLSFTVLGRGHSHTFENIRTSNGAALSAAADGL 120
Qy 121 FTIRGPKELSPSCNSLAVLPAATTNGKSQTPTTSTPSNGTIYSKTDLLLNNEKPSF 180
Db 121 FTIRGPKELSPSCNSLAVLPAATTNGKSQTPTTSTPSNGTIYSKTDLLLNNEKPSF 180
Qy 181 YSNLVSGDGGAIADAKSLTVQGISKLCVFQENTAQDGGACQVTSFSAMANEAPAFIAN 240
Db 181 YSNLVSGDGGAIADAKSLTVQGISKLCVFQENTAQDGGACQVTSFSAMANEAPAFIAN 240
Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Qy 301 NGKTLFLNNVSPVYIAAKQPTSGQASNTNNYDGGGAIKFCXGQAG-AGNNNGSVSPD 359
Db 301 NGKTLFLNNVSPVYIAAEQPTNGQASNTSDNYDGGGAIKFCXGQAG-AGNNNGSVSPD 360
Qy 360 GEGVVPFSSNAAGKGGAIYAKLSVANCGPVQFLNNTANDGGAIIYLGESGELSADYG 419
Db 361 GEGVVPFSSNAAGKGGAIYAKLSVANCGPVQFLNNTANDGGAIIYLGESGELSADYG 420
Qy 420 DIIIDGNLKRKTAKENAADVNGTVSSQAI SMGSGGKITTLRAKAGHQLLPNDPIEMANGN 479
Db 421 DIIIDGNLKRKTAKENAADVNGTVSSQAI SMGSGGKITTLRAKAGHQLLPNDPIEMANGN 480
Qy 480 NQPAQSSKLLKINDGEGTGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQGG 539
Db 481 NQPAQSSKLLKINDGEGTGDIVFANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQGG 540
Qy 540 SLYNEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 599
Db 541 SLYNEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 659
Db 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSNOKINVLKQLQGTQPSANAPSDITL 660
Qy 660 GNEPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGYNPGRVVASI VPSNLSMGSILDI 719
Db 661 GNEPKYGYQGSWKLAWDPNTANNPGYTLKATWTKTGYNPGRVVASI VPSNLSMGSILDI 720
720 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHRRDALGQGYRIYSGGYSLGANSYFGSSMF 779
721 RSAHSAIQASVDGRSVCRLWVSGVSNFFYHRRDALGQGYRIYSGGYSLGANSYFGSSMF 780
780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVYFGNQHMKT 839
781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSQOALCGSYLFGDAFIASVYFGNQHMKT 840
840 SYTFABESDVRWDNNCLVGEIGVGLPIVITPSKLYLNELRPFVQAEFSAHDESFTBEGD 899
841 SYTFABESDVRWDNNCLVGEIGVGLPIVITPSKLYLNELRPFVQAEFSAHDESFTBEGD 900
900 QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTQTLLSHQETW 959
901 QARAFKSGHLINLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTQTLLSHQETW 960
```



```
QY 281 DGNVAVGGIYSGYGNVAFLLNGKTLFLNNVAPVYIAAKOPTSQASNTSNYYDCGAI 340
Db 209 -----SNLCSGNV-----NPLFF-----TGNSATNGGAI 233
QY 341 FCNKQAQAGSNNSGVSVDGEGVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 400
Db 234 CCI--SDLNTSEKLSLACQETLFPASNAKEKGAIYAKHVLRYNGPVSPINNNAKI 291
QY 401 GGAIYLGSGELSLADYDIIIFDGNLKRKTAKENAAADVNTVSSQAISMGSGKITTLLR 460
Db 292 GCAIAIOGGSLILAGEGSLVFPNNQSORTSQGLVR-NAIYLEKDAI-----LSSLE 343
QY 461 AKAGHQIILFNDPI-EMANGNQPQAGSSKLLKINDGEGVTG-----DIVPA----- 504
Db 344 ARNG-DILFPDPIVQSSKESPLSPSLQASVTSPTPATAPLVQTSANRSVIFSSRL 402
QY 505 -----NGSSTLVQNTIEOGRVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
Db 403 SEBEKTPDNLSLOQOPIELKSGRLVLRKORAVLSAPLSQDPQALLINEAGTS----- 455
QY 556 PPQPPAANQLITLSNLHLSLSSLANNAVNTPTNPPAQDSHPAVIGSTTAGSVTISGP 615
Db 456 -----LKTSBDLKLATLSIPLHSL-----DTEKSVTIHAP 485
QY 616 -----IPFEDL-DDTAYDRYDMLGSKNQKINVLKQLGTCKPANAPSDLTL--GNEMPKY 666
Db 486 NLSIQIKIFLSNGSDENFYENVLLSKQN-NIPLLT-----SKESQSHLPLDGNLSHF 539
QY 667 GYQGSWKLAWDNTANNPPTLAKATWTKTGYNPGPERVASLVPNSLWSIILDIRSAI 726
Db 540 GYQGDWTSW--KDSDEG-HSLIANWTPKNVYHPEROSTLVANTLWNTYSDMQAVQSMI 596
QY 727 QASVDGRSYCRGLMWGVSNNFFY-HDRDALG-QGYRTISGGYSLGANSY-FCSSMFGIAP 783
Db 597 NTIAHGGAYLPQWGSVAVSNLFPYADHSGKPIDNWHRSLSGLVFGISTHSLDDHFSCLAA 656
QY 784 TEVFGSKDYVVCNRNHHACIGSVLSTQOALCGSYLFGDAFIRASGYFGNQHKTSY-T 842
Db 657 QOLLKSKSDSPITSTE-----TTSYIATVQALATPLM-KISAQACNESIHELKTKYRS 710
QY 843 PAESDVEDWNNCLAGIBAGLPIVITPSKLYNLRLRPFVQAEFSYADHESFTBEGDQAR 902
Db 711 FSKEGFGSHVAVSGEVACASIPVNSGSLP-SSFSIFSKLQGSFGTQDGFESSGEBIR 769
QY 903 AFKSGHLLNLSVGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHORTWTD 962
Db 770 SPSASSPRNLSLPMGSIITPEKKSQKTRNYFFLGAYIQDLKRDVBSGPPVLLKNVSWDAP 829
QY 963 AFHLARHGVVVRGSMYASLTSNIEVYHGRIYDRDASRGYLSAGSRVRP 1012
Db 830 MANLDSRAYMFLTNQORAL-HRLQTLNVSYYLRGQSHSYSLDLGTTYYR 878
```

RESULT 5

```
US-11-103-957-13
; Sequence 13, Application US/1103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; TITLE OF INVENTION: Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
```

```
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-13
```

```
Query Match 12.3%, Score 649.5; DB 11; Length 964;
Best Local Similarity 25.8%; Pred. No. 1.8e-39;
Matches 260; Conservative 155; Mismatches 389; Indels 205; Gaps 43;
```

```
QY 84 NLLGSPTVLGEGSLT---PENIR-----TSTNGAALSAAADGLFTIEGKELSPS 132
Db 39 SLSNKISLTGDTNHLTCYLDLNDRLYLAILOKTPEGAAV--TITDLSLFDTKQEGIFY 96
QY 133 NCNLSLAVLPAATTKGSOPTTTSNPTSGNTIYKTDLLLNNE--KFSFVSNLVSGD-- 188
Db 97 AKN-----LTPESGGAIGYASPNSTVEIRTI---GPVFNENTCCRLFTWRNYPADKI 149
QY 189 --GGAIDAKSLTVGIGISKLCVQENTAOAGGACQVVTFSAMANRAPIAPVANV----- 241
Db 150 REGGAIHAQNLINHNHDVVGMKNFSGVQGAISTANTFVVSQSCFLFMDNICIQTN 209
QY 242 -AGVGGGIAAVDQOGVSSSTSTEDPVVSPSNTAVEPDGNVAVVGGGIYS----- 293
Db 210 TAG-KGGAIIYA-----GTNSFE-----SNCDLFFINNACCAGGAIKPSICSLTG 254
QY 294 -YGNVAFLLNGKTLFLNNVAPVYIAAKOPTSQASNTSNYYGCGAIFCKNGQAAGSNN 352
Db 255 NRGNIVPNN--RCFKN-----VETASSEASDGGAIKVTTRLDVTGN- 294
QY 353 SGSVPDGEVVPFSSNVAAGKGAIIYAKKLSVANCGPVQFLRNIAND-GGAIYLGSGE 411
Db 295 -----RGRIFPSDNIKNYOGAIIYAPVVTLLVDNGPTVFINNIANNKGAIIYDGTN 346
QY 412 LSLSADYDIIIFDGNLKRKTAKENAAADVNTVSS-----QAISMGSGKITTLLRAKHQ 466
Db 347 SKISADRHAIIFENI-----VTNVTNANGTSTSANPPRRAITVASSSGEILLGAGSQN 402
QY 467 ILFPNDPEMANGNNQPAQSSKLLKINDGEGVTGDIVPAN--GSSTLYQN-----V 514
Db 403 LIFFDPLEVN-----AGVSVSFNKEADQTSVVSFGATVNSADFHQNLQTKTAPL 455
QY 515 TIBQGRIVLREKAKLSVNSLSQTGSLYMBAGS-----TWDFVTPOPPQPPAANQL 566
Db 456 TLSNGFLCIEDHAQLTVNRFTQTGGVVVSLGNAVLSCKYKNGTGD-----SAGNAS 505
QY 567 ITLSNLHLSLSSLIANNA-----VTNPTNPQAQDSHPAVIGST-TAGSVTISGPIPEDL 621
Db 506 ITLKHGLNLSSILKSGAEIPLLMVEPTN--NSNNYTADTAATPSLSDVKLS---LIDDY 560
QY 622 DDTAYDRYDMLGSKNQKINVLKQLGTKP--PANAPSDLTILGNB-----MPKYGYQGS 671
Db 561 GNSPYESTD-----LTHALSSQPMLSISEASDNQLOSENIDPSGLNVPHYGQQL 610
QY 672 WKLAW-----DPNTANNPPTLAKATWTKTGYNPGPERVASLVPNSLWS- 715
Db 611 WTMGWAKTQDPEPASSATITDPQKANRPHRTLLLTWLPAGVYVPSPKHRSPLIANTLMGNM 670
QY 716 ILDIRSAHSAIQASVDGRSYCRGLMWGVSNSFFPHDRDALCGGYRTISGGYSLG---ANS 772
Db 671 LLATESLKNSAELTPSGHPFW-GITGGGLGMVYQDPRENHPGPHMSSSGYSAGIAQQT 729
QY 773 YFGSMFGLAPTEVFGH-8KDYVVCNRNHHACIGSVLSTQQA-----LCSYSLFGDAFI 826
Db 730 HTFSLKEFSQYTKLNERYAKNNV--SSKNYSQCGEMLPSQGBGLLLTKLVGLSYGDHNC 787
QY 827 RASYFGNQHKTSYTPABESDVMDNNCIAGETAGLPIVITPSKLYNLRLRPFVQAEF 886
Db 788 HHFTYQG-ENLTSQGTFRSQT-----MGAVFFDLPMKPFGSTHILT--APFLGALG 836
QY 887 SYADHESFTBEGDQARAFKS-GHLLNLNLSVVPVGVKFDRCSSTH-PNKYSFMAAYICDAYRT 944
```

Db 837 IYSSLSHTEVGAYPRSTKTLPLNVLVPIGVKGSFNNATHRPQAWTVLAYQPVLYRQ 896
Qy 945 ISGTETTLSSHQETWTTDAFHARHGCVVVRGSMYAS-----LTSNIEVYG 989
Db 897 EPGIAAQLASKGIWFGSGSPSSRHAMSYKISQQTQPLSLWTLHFQYHG 945

RESULT 6
US-11-018-868-19
; Sequence 19, Application US/11018968
; Publication No. US20060034971A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Pincio, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-19

Query Match 12.3%; Score 649.5; DB 11; Length 964;
Best Local Similarity 25.8%; Pred. No. 1.8e-38; Indels 205; Gaps 43;
Matches 260; Conservative 155; Mismatches 389;

Qy 84 NLGSPVLGRGHSLT---FENIR-----TSTNGAALSNSAADGLFTIEGFKELSPS 132
Db 39 SLNKSISLGDTHNLNLCYLDNLRYLAILQKTPNEGA--TITDYLSPFDFTQKEGIYP 96
Qy 133 NCNLSLAVLPAATNKGSGQPTTTPSNGTIYSKTDLLLLNNE--KFSFYSLNLSGD-- 188
Db 97 AKN-----LTPESGGAIGYASPNSTPVEIRDITI---GPVIFENNTCCRLFTWRNPPYAADKI 149
Qy 189 ---GGATDAKSLTVQGISKLCVFQENTAQDGGACQVVTFSAMANEAPAFVANV----- 241
Db 150 REGGALHAQNLINHNHVVGVFMKFSYVQGGAI StantonFVVSNSQSCFLFMDNICIQTN 209
Qy 242 -AGVRGSGIAAVDQGGQGVSSSTSTEDPVVFSFRNTAVEPDGNVAVRGGIYS----- 293
Db 210 TAG-KGGAIIYA-----GTNSFE-----SNCDLFFINNACCAGGAIFSPICSLTG 254
Qy 294 -YGNVAFNLNGKTLFLANVASPVYIAAKOPTSQASNTSNYGDGGAIFCKNGAQAQGSNN 352
Db 255 NRGNIYFYNN--RCFKN-----VETASSEASDGGAIKVTTRLDVTGN- 294
Qy 353 SGVSFDEGEVGFSSNVAAGKGAIIYAKKLSVANCGPVQFLNIAND--CGAIIYLGESGE 411
Db 295 -----RGRIFFSDNITKNYGGAIYAPVITLVDNGPTTFINNANNKGGAIYIDGTSN 346
Qy 412 LSLSDYGDIIIPGNLKRKAENAAVDVGVTVSS-----QAISMGSGKITTILRAKAGHQ 466
Db 347 SKISADRHAIIFENI-----VTNVTNANGTSTSANPPRRAIITVASSSGEILLGASSQN 402
Qy 467 ILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVPAN---GSSTLYQN-----V 514
Db 403 LIYDPIEVSN-----AGVSVSFNKEADQTSVVFSGATVNSADPHQNLQTKTPAPL 455
Qy 515 TIEGRVILREKAKLSVNSLSQSGSLYMEAGS-----TWDVPTPQPQPQPPANQL 566
Db 456 TLSNGFLCIEDHAQLTNRFTQTGGVSLGNGAVLSCYKNGTGD-----SASNAS 505
Qy 567 IITLNLHLSSLSLANN-----VTNPPTNPPAQDSHPAVIGST--TAGSVTISGPIPEDL 621
Db 506 IITLKHGLNLSILKSABEPLLWVEPTN--NSNNYADTAATPSLSDVKLS---LIDDY 560

Qy 622 DDTAYDRYDMLGNSNOKINVLKQLQGTXP--PANAPSDLTJLGN-----MPKGYGQS 671
Db 561 GNSPYESTD-----LTHALSSQPMLSISEASDNQLOSENIDFSLNVPHYGMOGL 610
Qy 672 WKLAW-----DPTANNNGPYTLKATWTKGYNPGPERVASLVPNLSLWGS- 715
Db 611 WTWGAKTQDPEPASSATITDPQKANRPHRTLLLTWLPAGYVPSPKHSPLIANTLMGNM 670
Qy 716 ILDIRSAHSIAQASVDGRSYCRGLWVSGVSNFFYHRRDALQGGYRIISGGYSLG---ANS 772
Db 671 LLATESLKNASBELTSPSGHPFW-GITGGGLGMVYQDPRENHPGTHMRSSGYSAGIAQOT 729
Qy 773 YFGSSMFLAFTVEVGR-SKDYVVCVRNHHACISGVYLSLSTQQA-----LCSYSLFGDAFI 826
Db 730 HTFSLKFSQTYTKLNERVAKNV--SSKNYSQCEBMLFSLQEGFLLTKLVGLYSYGDHNC 787
Qy 827 RASVYFGNQHMKTSYTPABESDVRWNNCLAGEIGAGLPIVITPSKLYLNELRPPVQAEF 886
Db 788 HHFTYQG-ENLTSQGTFRSQT-----MGGAVFDDLPMPKPFSGSTHILT--APFLGALG 836
Qy 887 SYADHESFTEGDOARAFKS--GHLLNLSPVGVKFDRCSSSTH-PNKYSFMAAYICDAYRT 944
Db 837 IYSSLSHTEVGAYPRSPSTKTLPLNVLVPIGVKGSFNNATHRPQAWTVLAYQPVLYRQ 896
Qy 945 ISGTETTLSSHQETWTTDAFHARHGCVVVRGSMYAS-----LTSNIEVYG 989
Db 897 EPGIAAQLASKGIWFGSGSPSSRHAMSYKISQQTQPLSLWTLHFQYHG 945

RESULT 7

US-11-103-957-41
; Sequence 41, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1016
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-41

Query Match 11.7%; Score 615; DB 11; Length 1016;
Best Local Similarity 24.2%; Pred. No. 6.1e-36;
Matches 280; Conservative 163; Mismatches 402; Indels 312; Gaps 56;

Qy 3 TSPHKFELSLAYSCCSLNGGGYAAEIMVPOGIYDGETLTVSPFPYTVIGDPSGTVFSA 62
Db 8 TSP--CFLACLCISY-----GFSS---PQVL-----TPNVTTF-----KGGDVLYN 46
Qy 63 GELTLKXNL-----DNSIAALPLSCFQNLGSLFTVLGRGHSLSLTPENTRTNGAALSNSA- 116
Db 47 GDCAFNVVYAGAENGSI-----ISANGD---NLITQONHTLSP-----TDSQGPVLQNYAF 95
Qy 117 --ADGLFTTEGFKELSPS---NCN-----SLLAVLPAA 144
Db 96 ISAGETLTLKDFSLSPFSKNSVCGRKMISGKTVTSISGAGEVIFWDSNVGYSPLSIVPAS 155


```
QY 145 TTNGSQTP-----TTSTPSNGTIYSKTDLLLLLNNEKFSFYNLVSGDGAIDAK 195
Db 156 TPTPPAPAPAAASSLSPTVSDARKGSIFS-----ET 189
QY 196 SLTVQIGISKLVCFOBNTAQADGACQVVTSPSAMANEAPAFVANVAGVGGIAAVQDG 255
Db 190 SLEISGVKKGVNF-----DNNAGNFGTVFRGNSN-----NNAGSGSG-----227
QY 256 QGVSSSTSTEDPVVFSFRNTAVEPDGNVARGGIIYGNVAFNN-KGTLFLNVA-S 313
Db 228 -----SATTPTSFTVKCKG-KVSTPDNVASCQGVVYKGTFLPKDNEGGIFFRGNTAYD 280
QY 314 PVYIAAQPTSGQASNTNNYDGGAIKCK-NGAQAGSNNSGVSFPD-----GBG-----362
Db 281 DLGILA---ATSRDQNTET---GGGGVICSPDDSVKFEKNKGSIVFDYNPAKGRGGSILT 335
QY 363 -----VVFSSNVAAGKGAIYAKLSVA-NCGPVQFLRNIAANDGGAIIYL-----G 407
Db 336 KEPSLVADDSVVFNSNTAEKGGAIYAPTIDISTNGSILFERNRAEGGAICVSEASSG 395
QY 408 EGBELSLADYDGIIFDGNL--KRTAKENADV--NGVTSSQAIISMGGSGKITTLRAKA 463
Db 396 STGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVS-----LNASG 441
QY 464 GHQILFNDPIEMAN---GNNOPAQSSKLL--KINDGEGYTGDIVFA-----N 505
Db 442 LSKLIFYDPVQNNASAGASTPSPSSMPGAVTINOSGNSVIFTAESLTPSEKLOVLN 501
QY 506 GSSTLYQNVTIEQRIVLREKAKLSVNSLSQTSGLSVMAGSTWDFVTPQPPQPPAANQ 565
Db 502 STSNFPAGALTVSGGELVVTGATLTTCITATSGRVTLSGASLSAVA-----GAANN 554
QY 566 --LITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTIGSIPFEDLDD 623
Db 555 NYTCTVSKLGDLESFLTPNYKT-----AILGAD--GTVTVNSGSTLDLVM 599
QY 624 TAYDRYD---WLGSNQKINVLKQLGKTPPANAPSDLTGN-----EMPKYGVQGSWKL 674
Db 600 SEAEVYDNPLFVGS-----LTIPFVTLSSSSASNGVTNKSVTINDADAAHYQGSWSA 653
QY 675 AW-----DPNTANNPGVTLKATWTKTG---YNPGPERVASLVPSNLWGSIL 717
Db 654 DWTXPPLAPDAKGMVPTNNK---TLYLWTRPASNYGEYLDPOKGEVFNLSVAGS 709
QY 718 DIRSAHSAIQASVDGR--SYCRGLMVSG--VSNPFYHDDRDLGQGYRIISGGY-SLGANS 772
Db 710 ALRFTPTNGLKEHYVSRDGVFVSLHALGDYILNVTQDDRD---GFLARYGGFOATAASH 765
QY 773 YFGSMFLATEVFGRSKDYVVCESNKH-----ACIGSVYL---STQOALCGSYLFGDA 824
Db 766 YENGSIQFVAFQGLYGOTKSRMYYSKODAGNMTLSCFGRSYVDINGTETVN-----816
QY 825 FIRASYPGNQHKMTSY-----TPAEBEDVRWDDNCLAGETGAGLPV--LTPSKLYLN- 876
Db 817 YWETAYGYSVVRMTQYFNDTKQFDHCKCHWNHNYAFVGAENFLEYCIPRQFARD 876
QY 877 -ELRPPVQAFSADHESFTEEGQARAFKSHLLNLNSVPVGVKFDRCSSTH---PNKYS 932
Db 877 YELTGPMRFEMAGCWSSSTRETSGLTRYFARGSGHNSLPIGIVAHAVSHVRSPSPSKLT 936
QY 933 PMAAYICDAYTISGTETTLTSHOBTWTDFAFLAHGVVVRGSMYASLTSNIEVYGHGR 992
Db 937 LNMGYRDPIMRWVTPHCNMBEIIANGVKTPIQGSPLARH-----AFFLEVHDTLTHIFGR 990
QY 993 YEYRDASRGYGLSAGSR 1009
Db 991 -----AYMNYSLDARR 1002
```

RESULT 8

US-11-018-868-22

; Sequence 22, Application US/11018868

; Publication No. US20060034871A1

```
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (P23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1016
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-11-018-868-22
```

```
Query Match 11.7%, Score 615; DB 11; Length 1016;
Best Local Similarity 24.2%; Pred. No. 6.1e-36;
Matches 280; Conservative 163; Mismatches 402; Indels 312; Gaps 56;

QY 3 TSPHKFPLSMILAYSCCSLNGGGVAAEIMVPOGIYDGETLTVSPYTVIGDPSGTTVPSA 62
Db 8 TSP--CFIACLCYSY-----GPASS---PQVL---TPNVITPF-----KGDDVVLN 46
QY 63 GELTKLNL-----DNSIAALPLSCFNLGASFTVLGRGHSITFENIRSTTNGAALSNSA- 116
Db 47 GDCAFVNVAAGENSIS---ISANGD---NLITIGQHHTLSP---TDSQGPVLQNVAP 95
QY 117 --ADGLFTIEGFKELSPS---NCN-----SLAVLPAA 144
Db 96 ISAGETLTLKDFSLMFSKNVSCGEKMIKSTVSIISGAGEVIFWDSNVGVSPLSIVPAS 155
QY 145 TTNGSQTP-----TTSTPSNGTIYSKTDLLLLLNNEKFSFYNLVSGDGAIDAK 195
Db 156 TPTPPAPAPAAASSLSPTVSDARKGSIFS-----ET 189
QY 196 SLTVQIGISKLVCFOBNTAQADGACQVVTSPSAMANEAPAFVANVAGVGGIAAVQDG 255
Db 190 SLEISGVKKGVNF-----DNNAGNFGTVFRGNSN-----NNAGSGSG-----227
QY 256 QGVSSSTSTEDPVVFSFRNTAVEPDGNVARGGIIYGNVAFNN-KGTLFLNVA-S 313
Db 228 -----SATTPTSFTVKCKG-KVSTPDNVASCQGVVYKGTFLPKDNEGGIFFRGNTAYD 280
QY 314 PVYIAAQPTSGQASNTNNYDGGAIKCK-NGAQAGSNNSGVSFPD-----GBG-----362
Db 281 DLGILA---ATSRDQNTET---GGGGVICSPDDSVKFEKNKGSIVFDYNPAKGRGGSILT 335
QY 363 -----VVFSSNVAAGKGAIYAKLSVA-NCGPVQFLRNIAANDGGAIIYL-----G 407
Db 336 KEPSLVADDSVVFNSNTAEKGGAIYAPTIDISTNGSILFERNRAEGGAICVSEASSG 395
QY 408 EGBELSLADYDGIIFDGNL--KRTAKENADV--NGVTSSQAIISMGGSGKITTLRAKA 463
Db 396 STGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVS-----LNASG 441
QY 464 GHQILFNDPIEMAN---GNNOPAQSSKLL--KINDGEGYTGDIVFA-----N 505
Db 442 LSKLIFYDPVQNNASAGASTPSPSSMPGAVTINOSGNSVIFTAESLTPSEKLOVLN 501
QY 506 GSSTLYQNVTIEQRIVLREKAKLSVNSLSQTSGLSVMAGSTWDFVTPQPPQPPAANQ 565
Db 502 STSNFPAGALTVSGGELVVTGATLTTCITATSGRVTLSGASLSAVA-----GAANN 554
QY 566 --LITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTIGSIPFEDLDD 623
Db 555 NYTCTVSKLGDLESFLTPNYKT-----AILGAD--GTVTVNSGSTLDLVM 599
QY 624 TAYDRYD---WLGSNQKINVLKQLGKTPPANAPSDLTGN-----EMPKYGVQGSWKL 674
Db 600 SEAEVYDNPLFVGS-----LTIPFVTLSSSSASNGVTNKSVTINDADAAHYQGSWSA 653
```

```
QY 675 AW-----DPNTANNGPYTLKATWTKG-----YNPGRVASLVPNSLWGSIL 717
DB 654 DWTKEPLAPDAKGMVPPNTNN-----TLYLTWRPASNYGEYRLDPQKGLVPSNLWVAGS 709
QY 718 DIRSAHSAIQASVDGR--SYCRGLWVSG--VSNFFVHRDRDALGQGYRYISGGY-SLGANS 772
DB 710 ALRTTNGLKEHYVSRDVGFAVSLHALGDYILNYTQDDER-----GFLARYGGFOATAASH 765
QY 773 YFGSMFGLAPTEVFGRSKDYVYCRSNHH-----ACIGSVYL-----STQOALCGSVLFGDA 824
DB 766 YENGSI FGVAFGQLYGQTKSRMYSKDAGNMTMLSCFRYSYVDIKGTETVM-----816
QY 825 FRASYGFGNQHKTSY-----TPAESDVVRDNNCLAGEICAGLPV--ITPSKLYLN- 876
DB 817 YWETAYGYSVHRMHTQYFNDKTKQFDHDKCHWHNNNYAVFGAEHNFLEYCIPTRQFARD 876
QY 877 ELRPPVQAPESYADHESFTEBQDQARAFKSGHLNLSVPVGKFDRCSSTH---PNKYS 932
DB 877 YELTGPMREMAGCHSSSTRETGSLTRYFARGSGHNSLPIGIVAHAVSHVRSPPSKLT 936
QY 933 FMAAYICDAYRTISGTETLLSHOETWTTDAFHARHGVVVRGSMYASLTSNIEVYHGR 992
DB 937 LNWGYRDPDIWRVTPHCNMEIANGVKTPIQGSPLARH-----APFLEVHDTLYIHFGCR 990
QY 993 YEYRDSRGYGLSAGSR 1009
DB 991 -----AYMNYSLDARR 1002

RESULT 9
US-11-103-957-15
; Sequence 15, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
; US-11-103-957-15

Query Match 11.3%; Score 597.5; DB 11; Length 1531;
Best Local Similarity 24.1%; Pred. No. 2e-34;
Matches 293; Conservative 153; Mismatches 474; Indels 295; Gaps 54;

QY 17 SCCSLNGGGYAAETMV-----PQGI-YDGETLTVSPFYTVIGDPSTGVTFPSAGELTIXNL 70
DB 393 STACLGGAIAAQEIVSIQNNQAGISPEGGKASFGG-----GIACGSFSAGGASVLGTI 447
QY 71 D-NSIAALPLS-----CPGNLLGSTVLGRHSLTFENIRPTSTNGAALNSAADGLFTIEG 125
DB 448 DISKNLGAISFSRTLCTTSDIQGMEYQG-GGALFGENISLSENAGVLT-----494
QY 126 FKE---LSFNSCNLSLAVLPAATTKGSOPTTTSTPSNGTIYSKTDLLLLNNEKFSFYS 182
DB 495 FKONIVKTFASNGKILGGGAILATGKVEITNNSGISTGN--ARAPQALPTQEEPLFS 552
```

```
QY 183 N-----LVSG--DGAIDAASLTVQGI-SKLCVFOENTAQAD-----GGACQVVT 224
DB 553 KKEGRPLSSGSGGAILGREVALIH-NAAVFEQNRLOCSSEBATALIGCCGGGAVHGM 611
QY 225 SFSAMANEAPATAVANVA---GVRGGGIAAVODQOQGVSSSTSTEDPVPVSRNATVAFD 281
DB 612 STSIVGNSS-VRFGNNTAMGQGVSGGALLS-----KTVLAGNSGVDFPS 654
QY 282 GNVARVGGGIY--SYGNVAPLNNKXTL-----FLN 309
DB 655 RNIALSGGALQASEGNCCELVDNGVYLFPRDNRGVYGGAI-SCLRGDVVISGNKGRVFKD 714
QY 310 NVASPVYIAAK-----QPTSGQASNTSNY-----GD-----336
DB 715 NIATRLVYEETVEKVEEPAPEQKDNNELSFLGRAEQSFITAANQALFASSEDGDLSPES 774
QY 337 -----GGAIFCK-----NGAQAQSN-----SSSV-----SFQGE 361
DB 775 SISSEELAKRRECAAGAIPAKRVRIVDNQEAUVFSNNFSDIYGGAI FTGSLREEDKLDQ 834
QY 362 -----GWPFSSNVAAG-----GGAIYAKKLSVA-NCGPVQVFLRNINDGGA 403
DB 835 IPEVLISGNAGDVVPFSGNSSKRDHLPHTGGAICTQNLITISQNTGNVLFYNNVACSGGA 894
QY 404 IYLGESGELSLSADYDIIIPDGNLKRITAKENAADVNGVTVSSQAISM-GSGGKIITTLRAK 462
DB 895 VRIEDHGNVLLEAFGDDIVFKGNSSFRAQ-----GSDAIYFAGKESHITALNAT 943
QY 463 AGHQILFNDPIEMANGNNQPAOSSKLLKINDGE--GYTGDIVFANGSSTLYQNVITIEQR 520
DB 944 EGHAIVFHDALVFE--NLEERKSAEVLINSRENFGYTGSIINFLBAESKVPQCIHVQOQS 1001
QY 521 IVLREKAKLSVNSLSQ-TGGSILYMEAGSTWDFV-TPQPPQPPAAN-----QLITLSNLHL 574
DB 1002 LELLNGATLCSYGFQDQAGAKLVLAAGAKLKILDSTGPVQOQGHAI-SKPBATESSEPSG 1061
QY 575 SLSSLLANNAVTPNPTNPPAQDSH-----PAVIGSTTAGSVTIISG 614
DB 1062 AHSLSIAKNA-----QTTVPWVDIHTISVDLASFSSSQSGTVEAPQVI--VPGGSYVRSG 1115
QY 615 PIFFE--DLDDTAYDRYDWLGSNQKINVLKQLGTPKANAPASDLTLGN-----E 662
DB 1116 ELNLLELVNTGTGYENHALLKNEAKVPLMSFVASGDEASAEISLSVSDLIQHVVTPPEIB 1175
QY 663 MEKYGVQGSWKLAWDPNTANNPGYTLKATWTKGTGYNPGPERSVASLVPNSLM--GSIL--- 717
DB 1176 EDTYGHMGDWSEA-----KIQDG--TLVISWNPTGRLDPOKAGALVFNALWEGAVLSAL 1229
QY 718 -DIRSAHSAIQASVDGRSYCRGLWVSGVSNFFVHRDRDALG--QGYRYISGGYSLGAN--- 771
DB 1230 KVARPAHNLTAQORME-FDYSTNVWGFAPGFFRTLSENLAIDGYKAGYGGASAGVDIQL 1288
QY 772 -SYGSSMFGLAFTVEVFGRSK-DYVVVCRSNHHACTGSVTLSTQOALCGSVLFGDAFIRAS 829
DB 1289 MEDFVLGVSGAAFLGMDSQKFDAYSRK---GWVGSVYT-----GFLAGSWFFKQ 1337
QY 830 YGEG--NQHMKTSYTPAESDVVRDNNCLAGEICAGLPVITPSK--LYLNELRPPVQAE 885
DB 1338 YSLGETQNDMKTRYGVLGESSASWTSGVLADALVEYRSLVGPVRPTFFALHNPVFEVS 1397
QY 886 FSVADHESFTEBQDQARAFKSGHLNLSVPVGKFDRC--SSTHPNKYSFMAAYICDAYR 943
DB 1398 YASMKEPPTGTEQREARSFEDASLTNITIPLGKFKFELAFIKQFSEVNSLIGISYAEAYR 1457
QY 944 TISGTETLLSHOETWTTDAFHARHGVV-----RGSWYASLTNIEVYHGRVYERD 997
DB 1458 KVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTWSSYFSTVLGLTAF-CGGFTSTD 1516
QY 998 ASRGYGLSAGSRVF 1012
DB 1517 SKLGYEANTGLRLIF 1531
```



```
QY 58 TVFSAGELTLK--NLDNSIAALPLSCFGLNLLGSFTVLGRGH--SLTFENIRSTNGAALS 113
Db 80 NYIINRKLILSDPSLNVSS--GCAFRNLACKISFLGKNSSASHPFKHININGFCAGVF 137
QY 114 NSAADGLFTTIEGFKELSPNCSNLAVLPAATNKGSTPTTTTSPSNGTIYSKTDLLLL 173
Db 138 SSSS-----IEFTDLRKLAVF-----GSES-----TGGIFTAKEDISPK 171
QY 174 NNEKSPSYNLVSGDGAIDAKSLTVQGISK-----LCVFOENTAQAADGGACQV 223
Db 172 NNHHIAFRNITKNGGVQ-----LOGDMKGSVFDORGAIITNN-----QAV 217
QY 224 TSFSAMANEAPAFVANVAGVRGGGIAAVODQOQGVSSSTSTEDPVVFSRNTAVEFDGN 283
Db 218 TS--SSMKHSG-----RGAISGDFAGSR-----ILFLNNQOITFEGN 253
QY 284 VARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPT---SGOASNTSN---NYGDG 337
Db 254 SAVHGGAIYN-----KNGLVEFLGN--AGP--LAPKENTTIANGGAIYTSNFKANQOTS 303
QY 338 GAIFCKNGAQAGSNNSGVSF-----DREGVVPFSSNVAAGKGAIVAKKLSVANCGP 390
Db 304 PILFSONHA-----NKKGGAIYAQVYNLEQNDTIRFEKNTAKEGGAIYSSQCSITAHNT 359
QY 391 VOFLRNAND--GGAITYL--GESGELSLADYGDIIIPDGN--LKRTAKENAADVNGVTSS 445
Db 360 IIFSDNAAGDLGGAILLEGKPSLTLIAHSGNIAFSGNTMLHITKASLDHNSILIKE 419
QY 446 QAISMGGSGKITTLTRAKAGHQLFNDPIEMANGNNOQASSKLLKINDGEGYT-----G 499
Db 420 APYKI-----QLAANKNHSIHFFDPVMALSASSSPIQ-----INAPETPTPPSPKG 466
QY 500 DIVP-----ANGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTSGLSMEAG 546
Db 467 MIVFSGANLLDAREDVANRTSIFNPQVHLYNGTSLIENGHAHLIVOSFKQTGRISLSPG 526
QY 547 STWDFVTPO-----PPOQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSDHP 599
Db 527 SSLALYTMNSFFHGNISSEKPELEINGL--SPGVDISPSNLQAEIRAGNAPLR-----576
QY 600 AVIGSTTAGSVTISGP--IFFEDLDDTA--YDRYDWLGSNQKINVLKQLGTGKPPANAPS 655
Db 577 -----LSGSPSIHDPEGLFYENRDTAASPYQMEILLTSDKIVDISKF-----TT 620
QY 656 DLTGLNEMPKYGYQSGWKAWDNTANNGPY--TLKATWTKTG--YNGPDERVASLVNSLW 713
Db 621 DSLVTNK--QSGPQGAWHFSWQNTINNTKQILRASWLPTEGYVLESNRNVRGRAVPNSLW 678
QY 714 GSTLDIRSA-----HSAIQASVDGRSYCRGLWVGV-----SNPFYHDDRDLGQ 757
Db 679 STELLQTAHNLGDHLCNNRSLIPTSY--GVLIIGTGGAEMSTHSEESFISRLGATGT 737
QY 758 GYRYSIGVYSLGANSYFGSSMFLAPTEVFGRSKDVVCRSNHACIGSVY-----808
Db 738 SIIRLTPSLTSGG---GSMHFGDSP-----VADLPEHITSEGIQVONVGLTHVMG 784
QY 809 -LSTQOALCGSYLFGDAFIRASYFGNQHKMTSYTFABESDVWDNCLAGETGAGLPV 867
Db 785 PLTVNSLTCAA--LDHNAWVRIC-----SKDHTYG-----KWDTPGMRGTLGASVTF 831
QY 868 IPTSKLYNLRLPFVQAEFSYADHE-----SFTREGDQARAPKSGHLLNLVSPVGVKF 920
Db 832 -----EYDQTMR-----VPSFANIEATNIIQRAFTETGYNPRFSFKTKLLNIAIPIGY 881
QY 921 DRCSSTHPNYSFMA---AVICDAVRTISGTETTLISHQETWTD-----APHLARHG 970
Db 882 EFLCUN--SSFALLKGSIGYSRDIKENPSTLAHLAMNDFAMTNGCVPTSAHTLANQ 939
QY 971 VVVRGSMYASLTNSIEYHGRVEYRDASRGYGLSAGSRVRP 1012
Db 940 LILR---YKACSLYITAYTINR--EGKNLSN--SLSCGGYGVF 975
```

```
RESULT 12
US-11-018-868-15
; Sequence 15, Application US/11018868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00059 (PF23152.001)
; CURRENT APPLICATION NUMBER: US/11/018.868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 975
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-11-018-868-15

Query Match 10.5%; Score 551; DB 11; Length 975;
Best Local Similarity 24.9%; Pred. No. 2.4e-31;
Matches 279; Conservative 148; Mismatches 405; Indels 290; Gaps 57;

QY 3 TSFHKFFELSMILAYSCSLNGGYAAEIMV-----PQGIYDGETLTVSPPTVIGDPSCGT 57
Db 32 TLIPFLGALIVAPYSF-----AEMELAISGHKQG-KORDITFM-----ISSCPEGT 79
QY 58 TVFSAGELTLK--NLDNSIAALPLSCFGLNLLGSFTVLGRGH--SLTFENIRSTNGAALS 113
Db 80 NYIINRKLILSDPSLNVSS--GCAFRNLACKISFLGKNSSASHPFKHININGFCAGVF 137
QY 114 NSAADGLFTTIEGFKELSPNCSNLAVLPAATNKGSTPTTTTSPSNGTIYSKTDLLLL 173
Db 138 SSSS-----IEFTDLRKLAVF-----GSES-----TGGIFTAKEDISPK 171
QY 174 NNEKSPSYNLVSGDGAIDAKSLTVQGISK-----LCVFOENTAQAADGGACQV 223
Db 172 NNHHIAFRNITKNGGVQ-----LOGDMKGSVFDORGAIITNN-----QAV 217
QY 224 TSFSAMANEAPAFVANVAGVRGGGIAAVODQOQGVSSSTSTEDPVVFSRNTAVEFDGN 283
Db 218 TS--SSMKHSG-----RGAISGDFAGSR-----ILFLNNQOITFEGN 253
QY 284 VARVGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAKOPT---SGOASNTSN---NYGDG 337
Db 254 SAVHGGAIYN-----KNGLVEFLGN--AGP--LAPKENTTIANGGAIYTSNFKANQOTS 303
QY 338 GAIFCKNGAQAGSNNSGVSF-----DREGVVPFSSNVAAGKGAIVAKKLSVANCGP 390
Db 304 PILFSONHA-----NKKGGAIYAQVYNLEQNDTIRFEKNTAKEGGAIYSSQCSITAHNT 359
QY 391 VOFLRNAND--GGAITYL--GESGELSLADYGDIIIPDGN--LKRTAKENAADVNGVTSS 445
Db 360 IIFSDNAAGDLGGAILLEGKPSLTLIAHSGNIAFSGNTMLHITKASLDHNSILIKE 419
QY 446 QAISMGGSGKITTLTRAKAGHQLFNDPIEMANGNNOQASSKLLKINDGEGYT-----G 499
Db 420 APYKI-----QLAANKNHSIHFFDPVMALSASSSPIQ-----INAPETPTPPSPKG 466
QY 500 DIVP-----ANGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTSGLSMEAG 546
Db 467 MIVFSGANLLDAREDVANRTSIFNPQVHLYNGTSLIENGHAHLIVOSFKQTGRISLSPG 526
QY 547 STWDFVTPO-----PPOQPPAANQLITLSNLHLSLSLLANNAVTPNPPAQDSDHP 599
Db 527 SSLALYTMNSFFHGNISSEKPELEINGL--SPGVDISPSNLQAEIRAGNAPLR-----576
QY 600 AVIGSTTAGSVTISGP--IFFEDLDDTA--YDRYDWLGSNQKINVLKQLGTGKPPANAPS 655
Db 577 -----LSGSPSIHDPEGLFYENRDTAASPYQMEILLTSDKIVDISKF-----TT 620
```

```

QY 656 DLTGNEPKYQYQGSWKLAWDPNTANNPY-TLKATWTKG-YNPGERVASLVPSLW 713
DB 621 DSLVTNK--QSGFQAWHFSWQPNNTINNTKQILRASWLPTEGYVLESNRVGRVAPVPSLW 678
QY 714 GSILDIRSA-----HSAIQASVDGRSYCRGLWVSGV-----SNFFVHDRDALQ 757
DB 679 STFLLLQTASHNLGDLHCNRSLSLTPSYF-GVLIGTGAEMSTHSSEESFISRLGATGT 737
QY 758 GYRISGYSIGANSYFGSSMPLAFTEVFGRSKDYVVCNRNHHACIGSVY-----808
DB 738 SIIRLTPSLTLGG--GSHMFGDSF-----VADLPEHITSEGIQVNVGLTHVMG 784
QY 809 -LSTQOALCGSYLFGDAPIRASYFGNQHMTSYTFASESDVRDNNCLAGEIGAGLPV 867
DB 785 PLTVNSTLCAA-LDHNAMVRIC-----SKDHTYG-----KWDTFQMRGTLGASYTFL 831
QY 868 IPTSKLYLNELRPPVQAESFYADHE-----SFTBEGDQARAPKSGHLLNLSVPVGVKF 920
DB 832 -----EYDQWTR-----VPSFANIEATNQLQAFETETGYNPSFSKTKLLNIAPIGIGY 881
QY 921 DRCSTHPNKYSFMA-----AYICDAYRTISGTETTLSSHQBTWTD-----AFHLARHG 970
DB 882 EFCLGN--SSPALLGKSGISYSDIKRENPSSTLAHLAMNDPAWTTNGCSVPTSATLANQ 939
QY 971 VVRGSMYASLTSNIEVGHGREYDRASRGYGLSAGSRVFP 1012
DB 940 LILR---YKACSLYITAYTINR-EGKNLSN--SLSCGGYVGF 975

RESULT 13
US-11-018-868-50
; Sequence 50, Application US/1101868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Gullio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Finco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 50
; LENGTH: 975
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-50

Query Match 10.5%; Score 551; DB 11; Length 975;
Best Local Similarity 24.9%; Pred. No. 2.4e-31;
Matches 279; Conservative 148; Mismatches 405; Indels 290; Gaps 57;

QY 3 TSFHKFFLSMILAYSCSLNGGVAEIMV-----PQGIYDGETLTFSPVTVIGDPSTG 57
DB 32 TLIPKFLIGALIVVAPYSF-----AEMELAISGHKQG-KORDTFTM-----ISSCEPT 79
QY 58 TVPSAGELTLK--NLDNSIAALPLSCFNLGLSPFVLGRGH--SLTPENIRTSNGALS 113
DB 80 NYIINRKLISDFSLNKKVSS--GAFNRNLAKGISFLGKNSASIFPKHININGFGAGVF 137
QY 114 NSAADGLFTIRGKELSPNCSNLIALVLPAAATNKGSTPTTTSPPSNGTIYSKTDLILL 173
DB 138 SESS-----IEFTDLKLVAF-----GSES-----TCGIFTAKEDISPK 171
QY 174 NNEKPSFYSNLVSGDGAIDAKSLTVQISK-----LCVQENTAOAGGACQV 223
DB 172 NNHIAFARNITKNGGVQI-----LQDMKGSVSVFQDQGAIIFTNN-----QAV 217
QY 224 TSPSANEAFIAFVANVAGVGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGN 283

```

```

DB 218 TS-SMKHSG-----RGGAISGDFAGSR-----ILFLNQQITFEGN 253
QY 284 VARYGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAQPT---SQQASNTSN---NYGDG 337
DB 254 SAVHGGAIYN-----KNGLVEFLGN-AGP--LAPKENTTIANGGAIYTSNFRANQOTS 303
QY 338 GAIFCKGQAQAGSNNSGSVSF-----DGEVVVFPSSNVAAKGKGAIYAKKLVSANGCP 390
DB 304 PILFSQNH-----NKGGAIIYAQVNLNQODTIRFEKNTAKEGGGAISSQCSITAHNT 359
QY 391 VOPLRLAND--GGAIVL-GESGELSADVGDIIFDGN--LKRTAKENAADVNGVTYSS 445
DB 360 IIFSDNAGDLGGGAILLEGKPSLTLLAHSGNTAFSGNTMLHTTKKASLDHRHSILIKR 419
QY 446 QAISMGGGKITTLIRAKAGHOILENDPIEMANGNNQPAQSSKLLKINDGEYTT-----G 499
DB 420 APYKI-----QLAANKNHSIHFFDPVMALSASSPIQ-----INAPETETPPSPKG 466
QY 500 DIVP-----ANGSSTLYQNTTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAG 546
DB 467 MIVFSGANLLDAREDEVANRTSIFNQPVHLYNGTSLIENGHAHLIVQSPKQTGGRISLSPG 526
QY 547 STWDFVTPQ-----PPOQPPAANQLITLNLHLSSLANNANAVTNPTNPQAQDHP 599
DB 527 SSLALYTNWSPFHGNISKEPLEINGL--SPGVDISPSNLQAEIRAGNAPLR-----576
QY 600 AVIGSTTAGSVTISGP--IPFEDLDDTA--YDRYDWLGSNQKINVLKQLGKTKPPANAP8 655
DB 577 -----LSGSPSIHDPGLFVENRDTAASPQMEILLTSDKVIDISKF-----TT 620
QY 656 DLTGNEPKYQYQGSWKLAWDPNTANNPY-TLKATWTKG-YNPGERVASLVPSLW 713
DB 621 DSLVTNK--QSGFQAWHFSWQPNNTINNTKQILRASWLPTEGYVLESNRVGRVAPVPSLW 678
QY 714 GSILDIRSA-----HSAIQASVDGRSYCRGLWVSGV-----SNFFVHDRDALQ 757
DB 679 STFLLLQTASHNLGDLHCNRSLSLTPSYF-GVLIGTGAEMSTHSSEESFISRLGATGT 737
QY 758 GYRISGYSIGANSYFGSSMPLAFTEVFGRSKDYVVCNRNHHACIGSVY-----808
DB 738 SIIRLTPSLTLGG--GSHMFGDSF-----VADLPEHITSEGIQVNVGLTHVMG 784
QY 809 -LSTQOALCGSYLFGDAPIRASYFGNQHMTSYTFASESDVRDNNCLAGEIGAGLPV 867
DB 785 PLTVNSTLCAA-LDHNAMVRIC-----SKDHTYG-----KWDTFQMRGTLGASYTFL 831
QY 868 IPTSKLYLNELRPPVQAESFYADHE-----SFTBEGDQARAPKSGHLLNLSVPVGVKF 920
DB 832 -----EYDQWTR-----VPSFANIEATNQLQAFETETGYNPSFSKTKLLNIAPIGIGY 881
QY 921 DRCSTHPNKYSFMA-----AYICDAYRTISGTETTLSSHQBTWTD-----AFHLARHG 970
DB 882 EFCLGN--SSPALLGKSGISYSDIKRENPSSTLAHLAMNDPAWTTNGCSVPTSATLANQ 939
QY 971 VVRGSMYASLTSNIEVGHGREYDRASRGYGLSAGSRVFP 1012
DB 940 LILR---YKACSLYITAYTINR-EGKNLSN--SLSCGGYVGF 975

RESULT 14
US-11-013-957-11
; Sequence 11, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Yves
; APPLICANT: Lobet, Yves
; APPLICANT: Francois-Xavier Jacques
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12

```

```
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-11

Query Match          9.9%; Score 522.5; DB 11; Length 1034;
Best Local Similarity 23.6%; Pred. No. 2.9e-29;
Matches 274; Conservative 157; Mismatches 435; Indels 295; Gaps 53;

QY 14 LAYSCSLNGGGVAAIMVPOGIYDGETLTVSF-----PVTVIGDPSGTTVFSAGEL 65
D 7 LSFACLSF---FYLSTISILQA---NETDTLQFRFTFSDRLEIQVLDPA--SLITAQNI 58
QY 66 TLKNTLNSIAALPLSCFGLLGSFTVLGRGHSILTFENIRTSNGAALSNAADGL---FT 122
D 59 VLSNLOS-----NGTGACTISGNTQTQIFSN---SVNTTADSGGAFDMVTTST 104
QY 123 IEQFKELSFNSCNLSLAVLPAATNKGSGTPTTTSTPNSGTIYSKTDLLLNNEKFSFYS 182
D 105 ASDNANLLP--CNVY-----CTHNG-----GGAIRSGGPFRFLANNQDVLFPN 145
QY 183 NLVSG-----DGAIDAKSLTVQGISKLCVFOENTAOADGGACQVVTFSFA- 228
D 146 NISAGAKYVGTGDHNEKNRGALYATTITLTG-NRTLAFINNMSGDCGGAISADTQISIT 204
QY 229 -----MANEAPIAFVANVAGVGGIAAVQDGOOGVSSSTSTEDPVVPSRNT 276
D 205 DTVKGLIFENNHTLNHIPTQAENMA--RGAICSRD-----LCSISNNSGPV----- 252
QY 277 AVEFDGNVARVGGIYSYGNVAFNNKTLFLANNVAPVYIAAKQPTSGOASNTSNGYD 336
D 253 ---FNTNQGKGGAISATRCVIDNKKERIIFSN-----SSLGWSQSSASN 296
QY 337 GGAIFCKNGAAGNSNGSVSPGEGVVPFPSSNVAAGKGAIYAKKLSVANCGVQVPLRN 396
D 297 GGAIQTTQGTFL--RNKKGSIYFD-----SNTATHAGGAINCGVIDIRNDGPFVFLNN 347
QY 397 IANDGGAIVLGE--SGELSLADYGDIIIP-----DGNLKRKTAKENADVNGVTVSSQ 446
D 348 SAANGAAFLNLSKPRSATNYIHTGTGDIVFNNVVFLDGNL--LGKELPHINNNEITPY 405
QY 447 AISWGGGKITTILRAKAGHOILFND-----PIEMANGNPOAOSK-LKKNIDGEGYTD 500
D 406 TLSLG-----AKQDRIYFDLPQWERVKENTSNNPSPSTRNTIWNPEPESGA 456
QY 501 IVFA-NGSSTLYQNV-----TTEQGRIVLRKAKLSVNSLSOTGS---LYM 543
D 457 VVFSYNQMSDRIITLMGKEHNYIKEAPTTLKFGTLAIEDDAELEIFNPQNTPTSLAL 516
QY 544 EAGSTWDFVTPPPQPPAANQLITLSNLHLSLSSLLANNAVNTNP---TNPQAQSHPA 600
D 517 GSGATLV-----GRHGKUNITNLGVILPIILKEG--KSPPCIRVNPQDMTQNTG 564
QY 601 VIGSTTAGSVTISGPIPF-----EDLDDTAYDRYDWLGSNOKINVLKQLQGTKPPAN 652
D 565 T-GQTPSTSSISIPMIIFNGRLSIVDENYESVDSNDL--SRGKAQLILSIETNDGQ 621
QY 653 APSD--LTLGNEM---PKYGVQGSWKLAWDNT----- 680
D 622 LDSNQSSLNTSLSPHYGQGLWTPNWIITTTITILNNSAPTSAISAEOKTSET 681
QY 681 -----ANNG-----PYTLKATWTKTGYNPGPERSVSLVFN 710
D 682 FTPSNTTTTASIPNIKASAGSGSASNSGEVTTIKHTLVNWPVGVIVDPIRGDLIAN 741
```



```
Db 253 ---FNYNGGKGAISATFCVIDNKKERIIIFSN-----SSLGWSQSSASN 296
Qy 337 GGAIFCKNGAAGSNNSSVSPDGGVFFSSNVAAGKGAIYAKLSVANGCPVQFURN 396
Db 297 GGAIOQTQGFLL-RNNKGSIVPD-----SNTATHAGGAINCGYDIDRNGPVYFLNN 347
Qy 397 IANDCGAIYLCB--SGELSLADYGDIIIF-----DGNLKRTAKENAADVNGTVSSQ 446
Db 348 SAAMGAAPNLKSPRSATYIHTGTGDIVNNNVFTLDGNL--LCKRKLPHINNNEITPY 405
Qy 447 AISMGSGGKITTLRAKAGHQLFND-----PIEMANGNNQPAQSSK-LLKLDGSGYTD 500
Db 406 TSLSG-----AKKDTRIYFDLPOWERKENTSNPPSPTRNITVNPTEFPGA 456
Qy 501 IVFA-NGSSTLYQNV-----TIEQGRIVLRBKAKLSVNSLSOTGS---LYM 543
Db 457 VVFSYNQMSDITRLMGKEHNYKEAPTTLKFGTLAIEDDAELEIFNIPQTNPSTLLAL 516
Qy 544 EAGSTWDFVTPQPPQPPAANQLITLSNLHLSSLANNVTPNP-----THPPAQDSHPA 600
Db 517 GSGATLV-----GKHGKLNITWLGVLPIILKEG--KSPPCIRVNPQDMQTQNTG 564
Qy 601 VIGSTTAGSVTISGPIFF-----EDLDDTAYDRYDMLGNSQKINVLKQLGKPPAN 652
Db 565 T-GQTPSSTSSISTPMIIFNGRLSIVDENYSVDSMDL--SRGKAEQLILSIETNDGQ 621
Qy 653 APSD--LTLGNEM---PKYGYGSGMKLAWDPNT----- 680
Db 622 LDSNQSSLNTSLLSPHYGQGLWTPNWTITTYITILNNSAPTSATSIAEQKKTSET 681
Qy 681 -----ANNG-----PYTLKATWTKTGYNPGPERVASLVPN 710
Db 682 FTSPNTTASIPNIXASAGSGSASNSGEVTTIKHTLVVNAWPVGYTVDPIRRGLIAN 741
Qy 711 SLWGSILDIRSAHSAIQASVDGRSFCRGL-----WVS-----GVSNEFFYHDDRALG-QG 758
Db 742 SL-----VHS-----GRNTWGLRSLLPNSFWALQGAATLFTKQKGLSYHG 785
Qy 759 YRISGGYSLGANSYFGSMPFLAFTVFGRSKDYVVCRRNHACIGSVYLSLTOOALCGS 818
Db 786 YSSAKGYTVSSQA---SGAHGKPLLSFSSQSSDKWKEKTNRLSSRYLLS---ALCFE 839
Qy 819 YLFGD--AFI-RASVGFNGHMKTSYTPAESDSDVWDNCLAGEIGAGLPVITPS-KLY 874
Db 840 HPMFDRIALIGAAACNYGTHNMRSPYGTGKSGKGFHSTTL-----GASLRCELDRDSMPLR 895
Qy 875 LNELRPFVQABFSYADHESFTEBGGQARAF--KSGHLNLNLSVPVGVKFDRCSSTHPN-KY 931
Db 896 SIMLTPFAQALFRTPEPASIRESGDLARLFTLEQHTAVS--PIGIKAYSDDTWPTLSW 954
Qy 932 SFMAAYICDAYRTISGTETLLSHQETWTTDAPHLARGVVVVRGSMYASLTSNIEVYGHG 991
Db 955 EMELAYOPTYMKREPLNTLIQNGSVNTTTPLAKHSFYGRGS-HSLKFSHLKLPANY 1013
Qy 992 RYERDASRGYGLSAGSRVF 1012
Db 1014 QAEVATSTVSHYNAGGALVF 1034
```

RESULT 16

```
US-11-103-957-45
; Sequence 45, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT FILING DATE: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
```

```
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1751
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-45
```

Query Match 9.0%; Score 475.5; DB 11; Length 1751;

Best Local Similarity 23.6%; Pred. No. 1.5e-25;

Matches 269; Conservative 148; Mismatches 454; Indels 269; Gaps 51;

```
Qy 35 GIYGEITLTVSPYTVVIGDPGSGTTFVSAGELTILKLNLSIAALPLSCFNNLLGSFTVLGR 94
Db 719 GIYAKKAKMSRIDQLNISSENSATEI--GGGICCKESLELDALVSLVTENLVGK---EGG 773
Qy 95 GHSLTFENIRTSWTGAALSNSAADGLFTIEGFKELSFSCNCSLLAVLPAATTNKGSOPT 154
Db 774 GLHAKTVNINSLKSGFSNNKANSST-----GVATTASAPAAAASLQAAAAVPS 826
Qy 155 TTSTPSSNGTIYSKTDLLLNNEKFPYSNLVSGDGGDAIDAKSLTVQGISKLCVFQENTA- 213
Db 827 SPATPT-----YSGVV---GGAIYGEKVTPSCQSGTCQFSGNOAI 863
Qy 214 -----QADGACQCVTVSFSMAANEAPAFV-----ANVAGVRGGGIAAVQD 254
Db 864 DNNPQSLSNLVQGGAIYAKTSLSIGSSDAGTSYIFSGNSVSTGKSQTTGQIAGG----- 917
Qy 255 GQQCVSSSTSTEDPVPVFSRNTA-----VEPDGNVAR-----VGGGIYSYGNVAPLNN 302
Db 918 ---AIYSPTVTLNCPATFSNNTASMATPKTSSEDSGSGNSIKDTIGAI---AGTATLS 971
Qy 303 GKTLPLNNVA---SPVYIAAKQPTSGQASNTSNNYGD-----GGAIFCKNGA-QAGSNN 352
Db 972 GVSFRSGNTADLGAAGITLANANTPSATGSONSITEKITLENGSFIFERNOAKRGAIY 1031
Qy 353 SGSVSPGEGVGVVPSNVAAGKGAIY-AKLSVANCGPVQPL----- 394
Db 1032 SPVSIXGNNT- PNQNTSTHDGSAIYFTKDATIESLGSVLFTGNVNTATQASATSQON 1090
Qy 395 RNIAIDCGAIYLGSGE-----LSLSADYGDIIIFDGNLKRTAKENAADVNGTVSS 445
Db 1091 TNYANYGAAP-GDPGTTQSSQTDAILTLASSGNITFSNN---SLQNNQGD-----TPAS 1142
Qy 446 QAISMGSGGKITTLRAKAGHQLFNDPIEMAN---GNNQPAQSSKKLKINDGEG---YTG 499
Db 1143 KFCISAGVVKL-SLQAAAGKTIISFFDCVHTSTKKIGSTQNVYET--LDINKENSNPYTG 1199
Qy 500 DIVPAN-----GSSTLYQNVITIEQGRIVLRBKAKLSVNSLSOTGS-LYMEAGSTWDFVTP 554
Db 1200 TIVFSSLHENKSYIPQWAILHNGTLVKEKTELHVVSFEQKEGSKLIMKPGAVL----- 1254
Qy 555 QPPQPPAANQLITLSNLHLSSLANNNA--VTPN----- 590
Db 1255 ---SNQNIANGALVINGLTLDLSMWGTPQAGEIPSPPELRIVATTSASGSGSVSSIPT 1311
Qy 591 NP-----PAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGNSQKINVLKQL 645
Db 1312 NPKRISAAAPSGSAATPTWSENKVFJTGDLTLIDPNGNFYQN-PLMGSDLDLVLKJL-- 1368
Qy 646 GTEKPPANAPS-----DLTL-GNEMPKYQYQGSWKLAMPNTANNPGPYTLKATWTKGNP- 699
Db 1369 ---PTNTSDVQVYDVLTLSGDLFPQKGYMGMTWTLDSNFPQTK-----LQARWTFDYRRW 1419
Qy 700 -----GPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLGVSGV 744
Db 1420 VYIPRDNHFNANSILSGNSHVVYKQGLINNMLN-----NARFDDIAY-NNFWVSGV 1470
```

```
QY 745 SNFFYHDDRDLGQGYRISGGYSLGANSYFGSS--MFLAFTVFGRSKDYVVCRSNHHAC 803
Db 1471 GTFLAQOQTPLSEESYYSRGTSVAIDAKPRQDFILGAAPSKWGTK--AIKWHNYFH 1528
QY 804 IGSVYLSLQQALCGSYLFGDAFIR-----ASYGFGNQHMKTSY--TFAEES 847
Db 1529 KGSEYSYQASVYGGKFLY---PLLKQHGKVALPFLIQGVVSYGHKIKHDDTTLYPSIHERN 1585
QY 848 DVRWNNCLAGIEGAGLPVITTPSKLYLNELAPFPVQAEFSYADHESFTEEGDQARAFKSG 907
Db 1586 KGDWEDLGLADLRISMDLK-BPSKSSKRITVYGELEYSSIRQKQFTEIDYDPRHFDCC 1644
QY 908 HLLNLSPVGVKFD-----RCSSTHNPKYSFMAAYICDAVRTISGTETILLSHQETWTTDA 963
Db 1645 AYRNLPLPVGCAVEGAINMNCNLTMYNKLAL--AYMPSIYRNNPVCKYRVLSSNEAGQVIC 1702
QY 964 PHLARHGVVVRGSMYASLTNSI-----EYVGH-----GRYEYRDASRGYGLSAGSRVRF 1012
Db 1703 -----GVPTRTSARAEYSTQLYLGPFWTLYGNYTIDVGMYTLSQWT-----SCGARMIF 1751

RESULT 17
US-11-018-868-16
; Sequence 16, Application US/11018868
; Publication No. US20060034871A1
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; APPLICANT: Grandi, Guido
; APPLICANT: Ratti, Giulio
; APPLICANT: Bonci, Alessandro
; APPLICANT: Fanco, Oretta
; TITLE OF INVENTION: Immunogenic Compositions for Chlamydia trachomatis
; FILE REFERENCE: 002441.00099 (PP23152.001)
; CURRENT APPLICATION NUMBER: US/11/018,868
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 16
; LENGTH: 1751
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-018-868-16

Query Match 9.0%; Score 475.5; DB 11; Length 1751;
Best Local Similarity 23.6%; Pred. No. 1.5e-25;
Matches: 269; Conservative 148; Mismatches 454; Indels 269; Gaps 51;

QY 35 GIYDGETLTVSPPTVIGDPSGTTVFSAGELTLKLDNSTAALPLSCFGNLLGSFTVLGR 94
Db 719 GIYAKKAKMSRIDQLMISENSATEI--GGGICCKESLELDALVLSLSTVENLGVK---EGG 773

QY 95 GHSLTPEINTRTSTNGAALSNSAADGLFTIEGFKELSPSNCNSLLAVLPAATTNKGSTQPT 154
Db 774 GLHAKTWNISLXSGFSFNKANSST-----GVATTASAPAAAASLQRAAAPS 826

QY 155 TTSTPSNGTIYSKTDLLILLANKEFYSYNLVSDGGDAIDAKSITVQGISKLQVQENTA- 213
Db 827 SPATPT-----YSGVV--CGAIYGEKVTFSCSGTCQFSGNQAI 863

QY 214 -----QADGACQVTVSPSMAANEAPFV-----ANVAVRGGGIAAQQD 254
Db 864 DNNPSQSLNVQGAAYAKTSLSIGSSDAGTSYFSGNSVSTGKSQTTGQIAGG----- 917

QY 255 GQGVSSSTEDPVPVFSRNTA-----VEPDGNVAR-----VGGGIYSYGNVAPLNN 302
Db 918 ---AIVSPVTLNCPATPSNNTASMATPKTSSDEGSGSGNSIKDTIGAI---AGTALTLS 971

QY 303 GKTLFLNNVA---SPVYIAAKQPTSGQASNTSNNYGD-----GGAIFCKXGA-QAGSNN 352
Db 972 GVSRFSGNTADLGAAGTLANANTPSATSGSQNSITEKITLENGSFIFERNQAMKRGAIY 1031

QY 353 SGSVSFPGEGVGVFPSSVNAAGKGAIY-AKKLSVANCGPVQFL----- 394
```

```
Db 1032 SPSVSIKNNIT-FNQNTSTHDSGAIYFTKDATIESLGSVLFTGNNTVATQASSATSQON 1090
QY 395 RNIAIDGGAIVLGRSSE-----LSLSADYGDIIIPDGNLKRKTAKENAADVNGVTVSS 445
Db 1091 TNTANYGAAP-GPPTQSSQTDAILTLASSGNITFSNN---SLQNNQGD---TPAS 1142
QY 446 QAINSGSGKITTLRAKAGHOILFNDPIEMAN---GNNQPAOSSKLLKINDGEG---YTG 499
Db 1143 KFCSIAGVVKL-SLQAAKGKTIISFFDCVHTSTKKGSTQNVYET--LDINKENSNPYTG 1199
QY 500 DIVPAN-----GSSTLYQNWITIEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTP 554
Db 1200 TIVFSSELHENKSYIPQNAILHNGTLVLKERTELHVVFSFEQKESKLIKMPGAVL----- 1254
QY 555 QPPQPPAANQLITLSNLHLSLSSLLANNA--VTNP-----PT 590
Db 1255 ---SNQNTANGALVINGLUTIDLSMGTQAGEISFPPELRIVATVTSASGSGSVSSIPT 1311
QY 591 NP-----PAQDHPAVIGSTTAGSVTISGPFIFFEDLDDTDAYDRYDMLGNSQNKINVLKQL 645
Db 1312 NPKRISAAAAPSGSRAATTPTMSENKVFLLGDLTLIDPNGNFYQN-PMLGSLDLVPLIKL-- 1368
QY 646 GTKPPANAPS-----DLTL-GNEMPKYIGQGSWKLAWDPNTANNPGYTLKATYTKTYN- 699
Db 1369 ----PTNTSDVQVYDLTSLGDLFPQKGYMGWTWLTDSNFTQGR-----LQARWTFDTYRM 1419
QY 700 -----GPERVASLVPNSLWGSTIDIRSAHSAIQAASVDGRSYCRGLWVSGV 744
Db 1420 VYIPRDNHFIANSILGSONSMIVVQGLINMLN-----NARFDDIAY--NFWVSGV 1470
QY 745 SNFFYHDDRDLGQGYRISGGYSLGANSYFGSS--MFLAFTVFGRSKDYVVCRSNHHAC 803
Db 1471 GTFLAQOQTPLSEESYYSRGTSVAIDAKPRQDFILGAAPSKWGTK--AIKWHNYFH 1528
QY 804 IGSVYLSLQQALCGSYLFGDAFIR-----ASYGFGNQHMKTSY--TFAEES 847
Db 1529 KGSEYSYQASVYGGKFLY---PLLKQHGKVALPFLIQGVVSYGHKIKHDDTTLYPSIHERN 1585
QY 848 DVRWNNCLAGIEGAGLPVITTPSKLYLNELAPFPVQAEFSYADHESFTEEGDQARAFKSG 907
Db 1586 KGDWEDLGLADLRISMDLK-BPSKSSKRITVYGELEYSSIRQKQFTEIDYDPRHFDCC 1644
QY 908 HLLNLSPVGVKFD-----RCSSTHNPKYSFMAAYICDAVRTISGTETILLSHQETWTTDA 963
Db 1645 AYRNLPLPVGCAVEGAINMNCNLTMYNKLAL--AYMPSIYRNNPVCKYRVLSSNEAGQVIC 1702
QY 964 PHLARHGVVVRGSMYASLTNSI-----EYVGH-----GRYEYRDASRGYGLSAGSRVRF 1012
Db 1703 -----GVPTRTSARAEYSTQLYLGPFWTLYGNYTIDVGMYTLSQWT-----SCGARMIF 1751
```

RESULT 18

```
US-11-103-957-21
; Sequence 21, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
```


Db 1080 SSSGQTDNLPLKLTASGNCIFRNNEYPSTSDTGTSTFCSIAGDVKLTMQAAKGKIS 1139
Qy 469 FNDPIEMA-NGNQPAQSKLLKINDGE-----GYTGDIVPAN-----GSTLYQNTYI 516
Db 1140 PFDALRTSTKKTGTQATAYDTLIDINKSEDSSETVNSAFTGTILFSSSELHENKSYIPQNVVL 1199
Qy 517 EGGRIVLEKAKLSVNSLSQ--TGGSLYMEAGSTWDFVTPQPQPPAANOLITLSNLHLS 575
Db 1200 HSGSLVLPENTELHVISFEQEGSSLVMTPGVLSNQT-----VADGALVINNMID 1251
Qy 576 LSSL-----LANNAVTNPT-----NPPAQDSH----- 598
Db 1252 LSSVEKNGIAEGNIFTPPELRIIDITTCGSGGTPSTDSESNQNSDDTEONNNDASNOQE 1311
Qy 599 -----PAVIG-----STTAGSVTISGPIFPEDLDTDAYD 627
Db 1312 SANGSSSPAVAAAHSTRNPAATAATPTTTTATTTNSQVILGGEIKLIDPNTFPQ 1371
Qy 628 RYDMLGSKNOKINVLKQLGTGPPANAPSLDITLGNEMPKYGYQSGWKLAWDPTANNPVT 687
Db 1372 N-PALRSDDQISLLVLPDSS--KMQAQKIVLTGDIAPQKGYTGT--LTLDPPQLQNG--T 1425
Qy 688 LKATWTKTGNPCGPVERVASLVP-----NSLWGSILDIRSAHSAL-----QASVDGRS 734
Db 1426 ISVLWKFDYSY-----RQAVYVRDNHIFYANSILGSQMLMVTYKQGLLNDKMLARFEVS 1480
Qy 735 YCRGLWVGVSFNFFHYDRDALOGVRYISGGVSLGANSYFGSSMP--GLAFTVEFGRSKDY 793
Db 1481 Y-NNLWISGLTMLSQVGTPTSEETYSRGSASVALDAPKPAHDVIVGAAFSRMIGKTKS- 1538
Qy 794 VVCRSNHHACISGVLSLTOALCGSVLFGDAFIR-----ASYGFGNQHKMKS 840
Db 1539 -LKRENNYTHKSEY--SYQASVYGGKPFHFVINKKTEKSLPLLQGVISYGIKHDVTTH 1596
Qy 841 Y-TFAEESDVRNNDNCLAGEIGA--GLPIVITPSKLYLNLRLPFPVQAEPFSYADHESFTEE 897
Db 1597 YFTIERNKGEWED---LGLWLTALRVSSVLVLPQAQDTRKTIYVGELEYSSIRQKQFTET 1653
Qy 898 GDOARAPKSGHLNLNSVPVGVKPD-RCSSTHNKYS-FMAAYICDAYRTISGTETTLHSH 955
Db 1654 EYDPRYFDNCTYRNLAIPNGLAPEGELSGNDILMNRFSVAVMLSIYRNSPTCKYQVLSS 1713
Qy 956 QETWTTDAFLARHGVVGRGSMYASLTSNI-----EVYCHGHEYEYRDASRGVLSAGSR 1009
Db 1714 GEGGEIIC-----GVPTNRSARGYSTQLYLGLPLWTLYGSYTIETADHTLAHMMNCGAR 1767
Qy 1010 VRF 1012
Db 1768 MTF 1770

RESULT 20
US-11-052-554A-1
; Sequence 1, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1268
; TYPE: PR1
; ORGANISM: Escherichia coli 0157:H7

US-11-052-554A-1
Query Match 4.5%; Score 237.5; DB 11; Length 1268;
Best Local Similarity 20.8%; Pred. No. 1.4e-08;
Matches 257; Conservative 155; Mismatches 421; Indels 401; Gaps 66;
Qy 12 MILAYSCCSLNGGGVAAEIMVPOG-IYDGETLTVSPYTVVIGDPSGTTVFSAGEITLKNL 70
Db 21 MMLSPFPVAS-----QAAGLVINKGTVYNNAN---GVPVVDINKPNSGL-SHNIWDLNVL 70
Qy 71 D-----NSTAALPLSCFGLLGSFTVLGRHSLTPEINIRISTTNGAALSNSAADGLFTI 123
Db 71 DKGVVFNNSANESTSLAGNIQGN-SNLTSGSAKVIILNEVTSKNPS-----TI 118
Qy 124 EGFKELSPSNCNSLIHLAVLPAATTNKGSQTPT-----TTSTP-----SNGTI-- 164
Db 119 NGMEVAGDKADLIILANPNGITVNGGSGINTGKLTLTGTPTDIODDKLAGYVNGGTITL 178
Qy 165 -----YSKTDLL-----LLNNEKFSYSLVSGDGAIDAISLTVQGISKLCYFQENTAGA 215
Db 179 GKLDNASPTILSRVVVNGKVSADELNVVAGN-NYVNAAGQVTVGSVSA-----TGSR 230
Qy 216 DGGACQVTVSPSAMANEAPLAFVANVAGVRCGGIAAVODGQGVGS----- 260
Db 231 NGYSVDVAKLGGMYANKISLVSTKGVGRNLGVIA--GGVNGVSDSKGNLLNSNAQIQ 288
Qy 261 -----SSTSTEDPVSFS-----RNTAVEF-DGNVARVG----- 288
Db 289 SASTNLTLTNGTLNNTGTVTSVGTISLNTNKTIVNTRAGNISITMGDIYVNSGTIDNTN 348
Qy 289 GGIYSYGNVAFLLNNGKTLPLNNVASPV-----YIAAKOPTSGOASNTNNGYDGGGAIFCK 343
Db 349 GKLAAGMLAVDTNNTATLINSKSGSSVGTIAGLVALKTGT-----LNNSNGQIRGGVVGLE 404
Qy 344 NGAQAGSNNSGVSPDGEVGFPSNVAAGK-----GAIYAKKL SVANGC 389
Db 405 SAAL--NNNGDIQTGTIAIISNGVNNKLIKRSSTGHIVIGAAGSVNNGSTKTADTG 462
Qy 390 PVQPL-----RNANDGGAIVLGESELSLSA-----DYGDIIIFDGN---LKR 429
Db 463 SDDSGLIATDQVEIGANNINNGQI--ASNGVSLSSYSTIDDDYAGKILNSKVIK 520
Qy 430 TAKENAADVNGVTVSSQAISSMGSGKITTILRAKAGHQLFNDPIEMANGNPOPAQSSKLL 489
Db 521 SSLEN--DTGGIS-GKQIEVAVGGSLT-----NNIGVISSE-- 554
Qy 490 KINDEGYTGD I-VFAN-----GSSTLYQNTYIEQ-----GRVLEKAKLSV-NSL 534
Db 555 -----EGDISLLANSVDNHHGFMGMQONITMESMSGVNNNTALIVASKKLLKINARGSI 606
Qy 535 SQTGGSLYMEAGSTWDFVTPQPQPPAANOLITLSNLHLSLSSLLANNA---VTNPPTN 591
Db 607 ENRDGNPF---GNAYGLYFGMPQ-----TCGMVKGKIELSGQNIYNNNSRLIAEDGPLT 659
Qy 592 PPAQ-----DSHPAVIGSTTAGSVTISGPIFPEDLDTDAYDRI--DWLGSNQKINVLKQ-- 644
Db 660 LQAKNTFDNTRALVTSAGDASIQVG-----TYNNYATVTSAGNLDIDATTLQNS 710
Qy 645 -LGTKEPPANA-----PSDLTLG-----NEMPKYI--QGSWKLAWDPTANNGP-YTLKAT 691
Db 711 SSGTMDNNATGFIASDKNLSLEVNSLNTYNGWISGKGDVDV-----TVNNGNLNRTI 765
Qy 692 WTKYTNPCGPVERVASLVPNSLWGSIL--DI-RSAHSAIOASVDGRSYCRGLWVGVSFNF 747
Db 766 AAEEKGLD-----IAALNGIENMKDISAGGDLTNTNTRHVTNNSNSNVGNQIVINAVNDI 820
Qy 748 -----PYHORD-----ALGQCYRYI-----SGVSLGANSYFGSMFGLAFTVEFGRSKDY 794
Db 821 NNRGNITVSADLNVTTKGNLYLYVMVGYGDIALSANSVANNATTIETATGLDIIDSKGNV 880
Qy 795 -VCRSNHHACISGV-----YLSLTOALCGSVLFGDAFIRASVYFGNQHM 837
Db 881 GNNRGNLHALNGVLSVKGNLNDNGEIRGYGDVTLATGNY---DSY----- 925

```
QY 838 KTSYTFABESDVRWNN-----CLAGE-----IGAGLPVITPSSKLYNELRPF 881
Db 926 KGSUT-SETGVDVTLTANIVDNAYGLIAGENSVDKASTIYNNTALIAANKKLVIN----- 979
QY 882 VQABFSYADHESFTEEGDQARAFKSGHLLNLNSVPVGVKFDRCSSSTHPNKYSFMAAYICDA 941
Db 980 AGGNLENRDGNFLR-----NNGALPGITDNG----- 1007
QY 942 YRTISGTEFTLLSHQETWTTDAFLARHG-----VVVRGSM----- 977
Db 1008 -GIVKGEVTLAQNVVNNSSIIAENGPLMLLSRGLDNRALLSGADAIIRAAGTF 1065
QY 978 ---YASLTS--NIEVY-----GHGRYEYRDAS 999
Db 1066 YNNYATTYSAGNLDVYAAASLANNASDGRLEDNTAT 1099

RESULT 21
US-11-052-554A-172
; Sequence 172, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 172
; LENGTH: 1643
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-11-052-554A-172

Query Match 4.4%; Score 232.5; DB 11; Length 1643;
Best Local Similarity 19.6%; Pred. No. 4.7e-08;
Matches 218; Conservative 161; Mismatches 437; Indels 297; Gaps 53;

QY 50 VIGDPSGTTVFSAGELTLKNDLSIAALPLSCFGLNLSFTVLGRGHSHTPENIRTS--- 106
Db 678 VAADPLNTNTTLADGTNLGSAENPLSTTHFATKAANADSILNVGKGNLYANNITTNDAN 737
QY 107 -----TNGAALSNSAADGLFTIEGPKELSPS-----NCHSLLAVL 141
Db 738 VGSLLHFRSGGTSIVSGTVGGQGHKLNNLILDNGTTVKFLGDITTFNGGKTKIEGSLQIS 797
QY 142 PAATNKGSGTPTTSTPSNGTI-YSKTDLLLNNEKFSYSN-----LVSGDGAII--- 192
Db 798 NNYTTH-----VESADNTGLEFVTDPTVTTLNKGCAFGLVKQVILISGPNIVNE 851
QY 193 DAKSLTVQGISKLVCFOENTAQA-----DGGACQVVTSPSAMAN-----EAPIAFVAN 240
Db 852 IGVGVIVHGIANSISFENASISGLTSLPLSGTPLDLVLIKTVGNGTVDNFAPVIVVSG 911
QY 241 VAGVRGG-----IAVQDQGVSSSTSTEDPVVPSRNT-----AVEFDGNVAVR 287
Db 912 IDSMINNGQIIIGKKNIIALSIGSD--NSITVNANTLYSGIRTTKNNQGTVTLSGGMNN 969
QY 288 GGGIYSYGNVAFNLNGK-----TLFLNVASPVYIAAKQPTSGQASN--TSNNYGDG 338
Db 970 PTIYGLG-----LENGSPKLKQVTTTDDYNLGSII-----ANNVTINDY----- 1010
QY 339 AIFCKNGAQAGSNNSGSV---SFDGEGVVPFSSN-----VAAGKGAIIYAKKLS 384
Db 1011 -VTLTGGIAGTDFDAKITLGSVNGNANVRVDSFTFSDPRSMIVATQANKGTVTVLGNAL 1069
```

```
QY 385 VANGC-----PVQPLRNIANDGGAILYGESGEL--SLSADYGDIIIPDGNLKRKTAKENAADVN 439
Db 1070 VSNIGSLDTPVASVRFTGNDSGA---GLQNIYISQIDFG-----TY 1108
QY 440 GVTVSSQAISMGSGGKITTLRAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGVTG 499
Db 1109 NLTLNSNVILGGG--TT-----AINGEIDL-----LTN 1135
QY 500 DIVPANGSSLYQNVTIEQ-----GRIVLEBKAKLSVNSLSQTGGSLYMEAGSTWD 550
Db 1136 NLIPANGTSTGWDNTSISTTLNVSSGNIGQVIAEDAQVNAATTTTITIKIDNANANF- 1194
QY 551 FVTPQPPQPPAAANQLITLSNLHLSSLSSLLANNAVTPNPPPAQDQSHPAVIGSTTAGSV 610
Db 1195 -----SGTOAYT-----L:QGGARFNGTIGAP-----NFAVTGSN----- 1224
QY 611 TISGPIF--PEDLDDTAYDRYDMLGSNOKINVLKQLGTGTPKANAP---SDLTIGNEMPK 665
Db 1225 -----IPVKYELIRDSNQD--YVLRTRNDVLNVVTTAVGNSAIAANAPGVQSNISRCLESTN 1278
QY 666 YGOGSKWLAWDPNTANN--GPYTLKATWTKTGN-PGPERSVASLVPSNLWGSI----- 716
Db 1279 TAAYNMMLAKDPDVATFVGAIAATDTSAAVTTVNLMDTKOTDLSNRL--GTLRLYLSNA 1337
QY 717 --LDIRSAHSAIQASVDGRSYCRGLWVGVSNNPFFYHDDALGQGYRYISGYSIGANSYP 774
Db 1338 ETSDVAGSATGAVSSGDEAEVSYGVMKPFYTIABQKGGIAGYKAKTGTGVVGLDTLA 1397
QY 775 GSSM-----FGLAFTEVFGRSKDYVVCRSNHHACIGSVILS---TQALCGSLYFGDAFI 826
Db 1398 SDNLMICAAIGITKDI--KHQDY---KKGDKTDINGLSFSLYGSQQLVKNPFAQGNAPF 1452
QY 827 RASYGONQMKTSYTFABESDVR-----WNNCLAGEIGAGLPVITPSSKLYNE 877
Db 1453 TL-----NKVKSQRYFFESNGKMSQIAAGNYDNNMTFGNLIFFGYDYNAMPNVL----- 1503
QY 878 LRPFVQAEFSYADHESFTEEGDQA--RAFKSGHLLNLNSVPVGVKFDRCSSST-----H 927
Db 1504 VTPMAGLSYLKSSNENYKGTGTTVANKRINSKFSRDLVLIVGAKV--AGSTVNTDIVIY 1561
QY 928 PNKYSFMA-----AYICDAYRTISGTETTLTLLSHQETWTTDAFLARHGTVVVRGSMYASLTS 983
Db 1562 PEIHSFVHVHVGKLSNSQSMLDQQTAPPISQPDRTAKTSVNI-----GLSANIKS 1612
QY 984 NIEV-YGHRYEYRDASRGYGLSAGS---RVRF 1012
Db 1613 DAKMEYGIG-YDFNSASK-YTAHQGTILKVRNF 1643

RESULT 22
US-11-052-554A-16
; Sequence 16, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2005-02-07
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-07-20
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 16
; LENGTH: 1250
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-16

Query Match 4.3%; Score 228; DB 11; Length 1250;
```


Qy 703 --RVASLVP-----NSLWSILDIRS-----AHSATQASVDG 732
Db 1276 VLNMAATLPLVFDLAELNSIRBLNIMKASPINNNVWGATYTRNNVTVDAGAFQTLTG 1335
Qy 733 -----RSYCRGLWVGVSFFVHRRDALGQGYRISGGYSIGA-NSYFGSSMFLGAF 783
Db 1336 MTVGIDSNDIPEGIATLGAFWGSHSHIGDRGHGVSIGYSLGGYASWHEHSGFYLDG 1395
Qy 784 TEVGRSKDYVVRCHNHACIGSVYLTQOALCGSYLFGDAFIRASYGFGNQHKMTSYTF 843
Db 1396 VVKLNRPSNVAGK-----MSGGAANGSY-----HSNGLGG-HIETGHRF 1435
Qy 844 AERSDVRWD 852
Db 1436 ---TDGNWN 1441

RESULT 24
US-11-087-099-1245
; Sequence 1245, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1245
; LENGTH: 3132
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-1245

Query Match 4.1%; Score 218; DB 11; Length 3132;
Best Local Similarity 20.6%; Pred. No. 1.3e-06;
Matches 228; Conservative 144; Mismatches 434; Indels 304; Gaps 51;
Qy 23 GGGYAAEIMVPGQIYD-----GETLTVSPFYTVIGDPGGTIV-FSAGELTLKN----- 69
Db 1035 GGGHVDVINAAGDVLGLMVSAGVSVIIDGGTGSILDNNGATANIANSASIKSDKAG 1094
Qy 70 -----LNSIAALPLSCFNLLGSFTVLGRGHSITPENIRT---STN-----GAALSNSAA 117
Db 1095 TASDALETTVDLAVEITGS--GKSFYIDESDALTSINAKVNNGSTNLNFTGGSFAFNAT 1152
Qy 118 DGLFTIEGFKELSFSCNCSLLAVLPAATTNKGSTPT-----T 155
Db 1153 TGAFSSSTGVGVTFENTGGVVAI-GTVTATGGSATITASTAITDATSITAIDTVVLNAGT 1211
Qy 156 TSTPSNGTIYKTDILLNNKFFSYNLVSGDGGDAIDAKSLTVQGISKLCVFOENTAQA 215
Db 1212 SIGASGSTIKSTATTLTLVASAGSIYAQ--SSDGATVNKAV---GTGKNIEPATTTGN- 1265
Qy 216 DGGACQVTSFSAMANEAPAFVANVAGVGGGIAVQDQGGVSSSTSTEDPVVPSRN 275
Db 1266 -----LTGTISAKGSVTLTATAGSVLSGG-----TSSSATGATVLSAG 1305
Qy 276 TAVRFDG-----NVARVGGIY--SYGNVAF-----LNNKTLFLNNVA 312
Db 1306 TAIGASGASVTVVAANLATSAGSVVLSNAGDVTLTAVATGAGFQLSNTGALVLNAGT 1365
Qy 313 SPVYIAAKQPTSGOASNTSNNGYDGAIFCKNGAQAQSNNSGVSFDCGEGVVFSSNVAA 372
Db 1366 -----AAQAVSLTAS-----GALTDCNGA---TNNISABSLTLVGLSIGST---- 1404
Qy 373 KGGAIIYAKKLSVANGCPVQLFNANDGGAIYIGE--SGELSLADYGDIIIFQNLKRT 430
Db 1405 ATNGAVDTQVNSV-----TATTTSGGIIYNELSGGLTITAATA-VGSDANVSIT 1453
Qy 431 AKENAADVNGTVSSQAISMGSGGKITTILRAKAGH-----QIL 468

Db 1454 GAGDIA-LGVITAKGDDVTLVSAGKITDNNGANNVTTADILNVTGPNQVGELETISITOLS 1512
Qy 469 FNDPIEMANGNOPAQSSKL-----LKINDGEGYTGDIVFANGSSTLYQNVVIE 517
Db 1513 STGSTDIVNAGAMAITKASLEGSSSPFAESLTLTDMAGDTATI--ANNISLTQLTTT-- 1568
Qy 518 QGRVLEKAKLSVNS--LSQTGSLYMEAGSTWDFTPQPPQPPAANQLITLSNLHLS 575
Db 1569 -GNIVF-----LDTNDTIVAQSGGTVTINAGTT-----DKGAVAIIGNITANONIS 1615
Qy 576 LSSLLANNAVTPPTPPAODSHPAVIGSTTA--GSVTISGPIFFEDLDDTAYDRYDWLG 633
Db 1616 ITA-----DSH-ITIGLGNAGIGDVSVD-XYVVD-----G 1646
Qy 634 SNOKINYL--KLQLGTKPPANAPSDLTLGNEMPK---YGYQGSWKLA-WDNTA----- 681
Db 1647 NSTAVNVIAARNFSLSGTTPTRQAEHLHTTNSIANAHADSEVAAKLTLLLEANTAAMDIMS 1706
Qy 682 ---NNGPYTLKATWTKTGYNPGERVASLVPNS-----LWGSILDIRSAHSAIQAS 729
Db 1707 TABSTANYSL--TLASSAYDTAQAEVDRDLAPIENGLKITTVLDGVSIALSTAASAIEVA 1764
Qy 730 VDGRSYCRGLWVSGVSNFFVYHRRDALCGGYRISGGYSILGAN-SYFGSSMFLGAFTEVFG 788
Db 1765 -----AGAAQAIPFGDAGSEAAAVTG---LAANVAGIAIAYALGIAHSEIAG 1809
Qy 789 RSKDYVVVCRNHHACIGSVYLSLTOALCGSYLFGDA--FIRASYGFGNQHKMTSYTFABE 846
Db 1810 QLDDAEDLAFNKAEMVAAKSTLTDLAIANHQAYKEATSISQAAVDAAVIERHNHMQVARQ 1869
Qy 847 SDVRWNNCLAGEIGAGLPIVITPSKLYLNELPFVQAEPSYADHESFTBEGDQARAFKS 906
Db 1870 AVTAEDASSTAGTLALPFGIQLS-GQLDVNAV-----NSDVVLAITG 1910
Qy 907 GHILAN-LSVPVGVKFCRCSSTHPNKYSFMAAYICDAVRTISGTETILLSHOETWTTDAFH 965
Db 1911 PAVLNSASATKTAGTGMSTATDNIISVGCALVADDIIRL---ETT----- 1953
Qy 966 LARHGTVVVRGSMYASLTSN--IEVYGHG 991
Db 1954 ---GQIINGNAGHSLTASLKFVGVAGTG 1977

RESULT 25
US-11-052-554A-90
; Sequence 90, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 90
; LENGTH: 3194
; TYPE: PRT
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-90

Query Match 4.1%; Score 218; DB 11; Length 3194;
Best Local Similarity 19.8%; Pred. No. 1.4e-06;
Matches 229; Conservative 126; Mismatches 403; Indels 396; Gaps 56;
Qy 14 LAYSCCSINGGGYAAETMWQGIYDGETLTVSPFYTVIGDPGSGTTFVSAGELTLK----- 68
Db 231 ITYNGSVNGNFGFDNVDNSGA-----TTI---SGVTFNNNGALTYKGGNGI 275

QY	69	-----NLDNSIAALPLSCFG-----NLLGSFTVLGRGHSLTENIRT	105
DB	276	GGSIITFTNSINHYKLNINANSVTFNNSTLGSMPNGNANTIGNAYIL--NANNITFNNL--	332
QY	106	STNGAALSNAADGLFTTEGFKELS-----FSCNCLLAVLPAAATNKGSTOPT	154
DB	333	TFNGGWFVFNRSDAHVNFGQTTINPTSPFVNMGTGTINPNALFNIQNTYPTIGNAYT	392
QY	155	-----TTSTPSN-----GTIYS---	166
DB	393	LFSMKNGNIAYDVNVLNWIIRLKNTOATKDNCKNATSNNTHTYVTVNGLGTLHYFRQ	452
QY	167	-----KTDLLLLNNEKFS-----FYSNLYSGDGAIDAKSLT-----VOGISKLCVFOBN	211
DB	453	IFSPDSIVLQSVYYGANNLYTNSVNIHNVFNLNKINDDRADTIFYLNGLNTWNYTOR	512
QY	212	TQAQDGACQVVTSTFSAMANEAPAFVANAVGRGGIAAQQDGGQGVSSSTSTEDPVVS	271
DB	513	FAQTYGGRKSALV-FNAITFPWANGALPKNSVTRFGYEGVWNGTGYITGTFADRVI	571
QY	272	FSRNTAVEPDGNVARVGGG-IYSGNVAFLNNGKTLF--LNNVASPVYIAAKQPTSGQAS	328
DB	572	-----IGNMMSGNAQTGGGATLNFVCGATEINIAGATFKNLKTTSONSYMTFMALNGSGS	627
QY	329	NTSN-----NYDGGALFCRNGAQAGSN--NSGSVSPDG--EGVVFSSNVAAG-----K	374
DB	628	GKINVSQSDFYDWTGCGYFTGNGVPSVNFNKAYYKFOGAENSYNFKNTNLAGNFKFQ	687
QY	375	GGAIAKLSVANCGVQF--LRNIAND-----GGAIYLGES-----GELSLS	415
DB	688	GKTTIEK--SVLNDASYAFDGVNNAFNEKFGGSPFNHAEQTNAPNNSFGSGSFSFN	745
QY	416	A---DYGDIIIFDG-----NLKRTAK-----ENAAVNGVTVS---SQAISMSGSGGKI	456
DB	746	AKQVDNGNSFGGVFNENNTPKASTNDTFNVNNQFKINGAQTDPTEKGVVFNMOGLL	805
QY	457	TLRAKAGHOILLPNDPIEMANGNQPAQSKLLKINDGEGYTGDIV-----	502
DB	806	SSLSVGTYYQLLNAKSVGYKDNNNALYQ---MLRWTSGENPSGKLDENKTPANSAKIYN	862
QY	503	-----PANGSSTLYQNVITTEQRIVLREKALSV-----NSLSQTGSSLYM	543
DB	863	VQPTDNGLYIYIKENFNNG-ITLUTRLCTGYTHCVNIDNDANFLKNVNNASNT--VFYL	919
QY	544	EAGSTWDFVTPQPPQPPAANQILITLSNLHLSLSLLANNAV-----NPPTNPPAQDS	597
DB	920	NGMTW-----KTAGTGYFTQD--YSGTNSVLVFNQTPPLAGANPTSN-----	961
QY	598	HPAIVIG-STTAGSV-----TISGPPIFEDLDDTAYDRYDVLGNSQKINVLKQLGTYKPPA	651
DB	962	--SVWFGKTSGAEWGLVGYIQGVFKANQIDIT-----GTIRSGNGAKTGGGATLVF	1011
QY	652	NAPSDLTGYN---EMPYKYGQSKLAWDPNTAN-----NGPYTLKA---TW	692
DB	1012	NAOERLIANANLNDKAGLQNSW-MNFIYVNGNLNVNANFNQTPPHGFNLUKANNITW	1070
QY	693	TKGTYNPGPBPAVSPNSLWGSITLIRSASAIQASVDGRSVCRLGWSGVSNFFYHDR	752
DB	1071	DK-----GS-----VSGGNGFVDNA	1086
QY	753	DALG-----QGYRYISGG---YSLGANSYFGSSMFLAFTVFGKSDYVVCRRNHACIG	805
DB	1087	NANGNAVIKNVNFSNDGTLIYKGGENS-----AGNSLTLENNTFNSY-----NINAKAQ	1135
QY	806	SVYLSSTOQALCGSVLFCD-----APIRASYGFGNQEM-----K	838
DB	1136	NLIFNNNSFNFGSGSYFNDTKRNVTPKGTNTILNSDPFSRLKGSVSDINNSIFNIERDUTDK	1195
QY	839	TSYTFABESDVRWDNNCLA-----GEIGAGLPIVITPSKLYLNLNLRPFVQ	883
DB	1196	TTTTLLSGDNIKYNNQALADNVFSKNLWDLIHVDEGEQST---LRTDNTTY-----FVQ	1246

QY	884	AEFSYADHESFTEE	897
DB	1247	--FTQSNQKQKVPB	1258

RESULT 26

US-11-052-554A-281

Sequence 281, Application US/11052554A

Publication No. US20050288866A1

GENERAL INFORMATION:

APPLICANT: Sachdeva, et al.

TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE

FILE REFERENCE: 30853/40359A

CURRENT APPLICATION NUMBER: US/11/052.554A

CURRENT FILING DATE: 2005-02-07

PRIOR APPLICATION NUMBER: US 60/589,227

PRIOR FILING DATE: 2004-07-20

PRIOR APPLICATION NUMBER: IN 173/DEL/2004

PRIOR FILING DATE: 2004-02-06

NUMBER OF SEQ ID NOS: 763

SOFTWARE: PatentIn version 3.3

SEQ ID NO 281

LENGTH: 5291

TYPE: PRT

ORGANISM: Escherichia coli 0157:H7

US-11-052-554A-281

Query Match 4.0%; Score 213; DB 11; Length 5291;

Best Local Similarity 21.8%; Pred. No. 6.5e-06;

Matches 182; Conservative 116; Mismatches 360; Indels 176; Gaps 33;

QY	39	GETLV-----SPPVTYIGDPSGTTVPFSAAGELT-LKNLDSIAALPLSCFGLN-----	85
DB	3267	QGLTLVTNLNGYQTTVTQDGSWSLTLPASDITALANNGYITATVTDLAGNLGSASKGV	3326
QY	86	-----IGSFTVLGRGHSLTFFENIRTS--TNGAALSNSAADGL-FTIEGFKELSFSCN	135
DB	3327	TYDVTAPVISFNTVAGDDVINNVHEHQAQIIISGTAGVAGDLVVTVIAGQQVVTSTAS	3386
QY	136	SLLAV-LPAATN---KGSQPTPTTTPSNGHYISKTDLLLLNNEKFSFYNLVSQDG--	189
DB	3387	GNWSGVPAFVSIGLADGTVTISATITDSAGNSSTQTHNVQVNTAAVLSVSTISGDNLI	3446
QY	190	GAIDAKS-LTVQGI--SKLCVFOENQAQDG-CACQVVTFSFSAMANEAPLAFVANVA-GVR	245
DB	3447	NAAEAGSALTLSGTGTNFATGTVTVLLNGKGYSATIQSGWSVNVNPAADVAALSDGTS	3506
QY	246	GGGIAAVDQCGQSVSSTSTEDPVVSFSRNTAVEFDGNVARVGGGIYSYGNVAFLNNGKT	305
DB	3507	YTVSASQAQDSAGNSSTQTHNVQV-----NTAAVLSVSTISGDNLIN-----	3550
QY	306	LFLNNVASPVYIAAKQPTSGQASNTSNNYDGGAI-FCRNG-----AQAGNSNGSVSFDG	360
DB	3551	-----AAEAGSALTLSGTGTNPATGTVTVLLNGKGYSATIQSGWSVNVNVP--	3597
QY	361	EGVVPFSSNVAKGGAIIYAKLSVANCGVPQFLRNIANDGGAIIYLGESGELSLSA--D	417
DB	3598	-----AADVAALSDGTSYTVSASAOQ-----SAGNSATASRSVAVDLTAPVTS	3640
QY	418	YGDIIIFDGNLKRKTAKENAADVNGVTVS--SQAISMSGSGGKIITTLRAKAGHOILLPNDPIEM	475
DB	3641	INTVSTDDRNLNAAEQOPLTLNGSTSAEVQTVTVTFGKTYTATVAANGTALNVPADV	3700
QY	476	ANGNPPAQSSKLLKINDGEGYTGDIVPANGSSTLYQNVITIE--QGRIVLREKAKLSVNS	533
DB	3701	LAALGQGAQTI-TASVNDRAGNPGQATHALTVDVTAPTVTIATVAGDDIINNAEQLAGQT	3759
QY	534	LSQT-----GGSLYMEAGSTWDFVTPQPPQPPAANQILITLSNLHLSLS	577
DB	3760	ISGTTTTAEVQGTVTVTFNGQTNATVSGSGSWSVFIP-----AQQFAGLSGSGYITS	3811
QY	578	SLLANNAVTPPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDRYDVLGNSNQK	637

382 RTNPIQAGTSAPVITGTAAALL--QKYPWMSN-----DNLRRT 420
QY 209 QENTAQADGACOVVTSF-----SAMANEPIAFVANAVRGCGGIAAQQDGOOGYS 260
Db 421 LLTTAQ-DIGAVGVDSKFGWGLLDAGKAMGPASPFPFGDTADTKG----- 465
QY 261 SSTSTEDPVVSPRNTAVFPGNVARVGGIYSYGNVAF-----LNNKGTFLNNVASPV 315
Db 466 -----TSDIAYSP-RNDISGTGGLIKGGSQLQHGNNTYTKTIIIEGGSILVYGNKSDM 520
QY 316 YIAAKQPT--SGQASNTSNYDGAIFCKNGAQAGSNNS-----GSVSPDGEVGFSSN 369
Db 521 RVETKALYNGAAGSGLN--SDGIVYLADTQSGANETVHIKGLQLDGKGLY----- 574
QY 370 VAAGK-----GGAIYAKK-----LSVANCG-PVQFLRNITADG 401
Db 575 TRLGKLLKVDGTAIIGGKLYMSARGKAGYLNSTGRRVPFLSAAKIGQDYSFPTNIETDG 634
QY 402 GAIYLGESGELSADYGD-----IIPDGNLKRITAK-----ENAADVNGVTVSSQA 447
Db 635 GLLASLDSVEKTAGSE-GDTLSYVVRGNARPTASAAHSAPAGLKHAVEQGSNLENLM 693
QY 448 ISMG-----SGKITTLLAKAGHQILFNDPIEMANGNNQPAOSSKLLKINDGEGYTGDIVFA 504
Db 694 VELDASESSATPETVETAA-----DRTDMPGIRPYGATFRAAAAVQHANAADGVRIEN 747
QY 505 NGSTLYQNVITB-----QGRIVUREKAKLSVNSLSOTGGSLYMEA-----GSTWD----- 550
Db 748 SLAATVYADTAHADMOGR-----RLKAVSDGLDHNGTGLRVIATQTDQDGGTWEQGV 802
QY 551 -----FVTPOPPQPPAANQLITLSNLHLSLSLLANNAVTPNPPAPQDQSHPA 600
Db 803 GWRGSGTQTVGIAAKGENTAA-----ATLGMGRSTWSENSE--NAKT-----DSISL 849
QY 601 VIG-STTAGSVTISGPFFEDLDOTAYDRYD-----WLGSNQ-----KINVLKQLGKTPP 650
Db 850 FAGIRHDAGDIGYKGLF-----SYGRYKNSISRSTGADHAEGSVNGTLQALGALG 902
QY 651 ANAP-----SOLITGN-----EMPKYGYOGSKLAWDPNTYANG-----PYT 687
Db 903 VNVFPAATGDLTVEGGLRYDLLQDAPAEKGS-ALGWSGNSLTEGTLVGLAGLKLQPLS 961
QY 688 LKATWTKYNGPGERVASLPVNSLWGSILDIRSA-----HSAIQASVDGRSYCR 737
Db 962 DKAVLFATA-----GVERDLNGRDYTVTGGTGATAATCKTGARNMPHRLVAGLADVEFG 1018
QY 738 GLWVSGVSNFFY-----HSDRALGQGYRYSIGYSILG-----ANSYFGSSMFLAFTV 786
Db 1019 NGW-NGLARYSYAGSKQYGNHSGRVGYRFLDGGGCTGSSDLANDSFIRQVLDROHFEF 1077
QY 787 FGR-----SKDYVVCRSNHHACIGSVYLSLTOQALCGSYLFGDAPIRASYGF-----GN 834
Db 1078 DGKYLFGSRGELAERSGH---IGLGIQSHQ--LGNLMIQQAIAKGNIGYVRFPSDGH 1132
QY 835 QHMTSTYFAEESD-----VRWDN-----NCLAGEIGAGLPIVITPSKL 873
Db 1133 EVHSPFDNHASHSDSDEAGSPVDFSLYRIHWDGYEHPHDPADYDGPQGGYPAPKGARDI 1192
QY 874 YLNLRLFPVO-AEFSYADHESFTTEGQOARAFKSGHLLNLNSVPVGVKFDRCSSHPNKYS 932
Db 1193 YSYDIKGAQVIRNLNDNRS--TQORLADRFHNAAGSMLTQGVGDPFK-----RATFYS 1244
QY 933 PMAAYICDAYRTISGTETTLTSHOETWTTDAFHLARHGVVVRGSMYASLTSNI-----EV 987
Db 1245 PELDRSGNAABFNGT-----ADIVKNIIGAAGEI 1274
QY 988 YGHRGYEYTRDASRGYGLSAGSRV 1010
Db 1275 VGAG-----DAVQ--GISEGSNI 1290

RESULT 29

US-11-052-554A-171

; Sequence 171, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 171
; LENGTH: 2340
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
; US-11-052-554A-171

Query Match 3.7%; Score 193.5; DB 11; Length 2340;

Best Local Similarity 19.6%; Pred. No. 5.1e-05; Indels 405; Gaps 45;
Matches 191; Conservative 109; Mismatch 270;

QY 55 SGTTPVS--AGBLTKQLDNSIAALPLSCFGLLGSFTVLGRHSILTFENIRTSNGAAL 112
Db 731 ASGTVLTDQTSBLTNN-DVVVNSNIITAGNNSKLIPTGNG-GIT-GNI--GANGAAL 785
QY 113 SNSAADGLFTIEG-FKELSFSCNLSLLAVLPAAATNKGSGTPTTTPSPNGTYSKTDLL 171
Db 786 QEVVFGTNTIGGTANSQNFVAHSAANVTITGLTTGALKYKDTGTITIAHGGILVGDID-- 843
QY 172 LLANKEFSYNLVSDGGGAIDAKSLTVQGISKLCYFQENTAQADGACOVVTSFSA-- 228
Db 844 -FNKAGP-----ILGDGAMIDGSLVLCNGGVAGTLDP-----IGDGNVTQNICADNANSI 893
QY 229 -----MANEAPIAFVANVAVRGGIAAQQDGOOGSVSSSTSTEDPVVVSFSR 274
Db 894 STINIQDNTKNTIAND--IFVDNIHFTNGGIL-----QLGGNLTTHN-----IDFGA 940
QY 275 N-TAVFDFGNVAVRGGIIYSYGNVAFNNKTLFLN-----NVA 312
Db 941 NGGTLEFNGN-----NTYNLNAIIVNGQNGILNAFTNLKASDDTIGTVKIIINIGQIG 992
QY 313 SP-----VYIAAKQPTSGQASNTSNYGD-----GGAIFC-- 342
Db 993 TPQNFYTIQVNNKXLTIVSSVNSINFGDANSQILSAPVDQTIKFNNLNETGGIITLD 1052
QY 343 -----KNGAQAGS--NNSGVSFDCG-----GVVFFSSNVA 371
Db 1053 SNGNLTISGNGIKLGSKNELSSLNIGKVTVTNDLDIQNTIHLNINNGALFDDQSLT 1112
QY 372 AGK-----GGAIY----- 379
Db 1113 SAKIKINICTVAGGATYTLDAINDFDLNTSGMVFQKQDSILLELNKSSNTNDHTITLTS 1172
QY 380 -----AKKLSVANCQPVQFLNIA-----DGAIVL----- 406
Db 1173 ALDPGNNGFGIILKLTDTNKLTIIDNNGNVAYTILGTANHLKQLTTFASIDNGALTKVGIN 1232
QY 407 -----GSGELSLSDYVDIIFDG 425
Db 1233 VENTLNIKDIELNEVNANVLFNKNTYTTATGNGINGHVDQGNAGVINLN--DDIIEIDG 1289
QY 426 NLKRTAKENAADVNGVTVSSQAISMGGKKIT-----TIRAKAGHQIL-----FNDPI 473
Db 1290 SVTST-----GNVNG-----TLNFNGSGKVGTGLINNI VMLQAGAGDVLSASNGYSITE 1338
QY 474 EMANGNNQ-----PAOSSKLLKINDGEGYTGDIVFANGSSTL----- 510
Db 1339 IQGNGNNNLTAANSHLTDTINKTGGQDLNLVFINGGSVSGSIGANAAVGDIIINAGSVN 1398


```
Db 1629 LLSPLGLHDFWQGYFNFSLNGYFVNSSFSFNATGGSFNFAVANSKIIFPENGDTID-----1684
Qy 80 SCFNGLLGSFTVLGRGHSITFENIRT--STNGAALSNSAAGLFTIEGFKELSPFNCSNL 137
Db 1685 --FSKYQCALIPASGVNSNI--NITTLNATNGLSL--NAGLNVSVQKGBICINLANC---1736
Qy 138 LAVLPAATYNGSQPTTT--TSTPSNGTIYSKTDLLLANNEKESFYSLNVSGDGGIDAOKS 196
Db 1737 -----PTTKNSSPANSVPTPTNESLSVHANNFTLGLTTIS--NGAIDLSQ 1779
Qy 197 LTVQGISKLCVFOEN--TAQDGGACQVTFVSFAMANEAPIAFVANVAGVGGGIAAVQDG 255
Db 1780 VTNSVIGTFLNENATLOANN-----LTIITAFNN-----1810
Qy 256 OQGVSSSTSTEDPVVFSFRTAVFEDGNVAVFGGIIYSYGNVAFNLNGKTLFLNNVASPV 315
Db 1811 ---ASNSTANIDGNFTLNQOATLSTNASGLNVGMNFNSYGLVF-----NLSHSV 1857
Qy 316 YIAAKQPTSGQASNTSNYDGGAIKFCNKAQAGSNSSGVDFDGBGVVFPSSNVAAGK 375
Db 1858 -----SHAIINTQGT-----ATIMANNPLIOFN-----ASSKE 1886
Qy 376 GAIYAKKLSVANGCPVQFLRNTANDGAIYLGESGELSADYDGIIFDGNLKRTA-----431
Db 1887 VGTY-----TLDSAKALIYYGNNOIT-----GGSSLDNLYLKALIDI 1925
Qy 432 --KENAADVNGVTSSQAI SMGSGGKITTLRAKAGHOI-----LFPDPIEMAN 477
Db 1926 NGKHWMTDNGLTYNQAVSVKDGGLVVGFKDSQNOYIVTSILYNKVIASVNDPI-----1981
Qy 478 GNNQPAQSSKLLKINDGEGYTGDIVFANGSSSTLYQWVTIEQGRIVLRKAKLSVNSLSQT 537
Db 1982 --NNPOA-----PTLKQYIAQIQ-----VOSVDSIDQA 2008
Qy 538 GGS-----LYMEAGSTWDFVTPQPPQPPAANOLITLNLHL---S 575
Db 2009 GGNQANLWLNKIPETKGSPLFAPYILSHSTKDLTT-----IAGDIANTLEIVANFPKND 2064
Qy 576 LSSLANNATNPPTNPQAQDSHPAVIGSTAGSVTISGPIFFEDLDOTAYDRYDWLGSN 635
Db 2065 ATNLIQNTYTQ-----QMSRLAKLSDTs-----IFARSDFLERLEALKKRFADAIPN 2113
Qy 636 QKINVLKLO-----LGTKPPANAPSD-----LTLGNEMPKYG 667
Db 2114 AMDVILKYSQRNRVQNNVWATGVGGASFISGGTGLYGINVGDYRFKGVIGG--YAAYG 2172
Qy 668 YQGSWKLAWDPNTANNGPYTLKATWTKGYNPGPERSVLPVNSLWGSILDIRSAHSAIQ 727
Db 2173 YSG-----PHANITQSGSNVNVGVYSRAFIKRSELTMSL--NETWYKTFPINSYDPL- 2224
Qy 728 ASVDRSRYCRLWVSGVS-----NFFYHRRDAL-----GQYRYI--SG-----764
Db 2225 LSIINGSYRYDTWTDAKINGYDFMFKDKSVIFKPOVGLSYIYIGLSGLRGIMDDPIYN 2284
Qy 765 -----GYSIGANSYFGSSMFLAFTVEFR 789
Db 2285 QPRANADPNKKSVLITINFALSRHYFNKNSYFYFIADV-GR 2324
```

RESULT 32

```
US-11-052-554A-179
; Sequence 179, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
```

```
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 179
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Shigella flexneri 2a str. 2457T
US-11-052-554A-179
```

```
Query Match 3.6%; Score 190; DB 11; Length 955;
Best Local Similarity 20.7%; Fred. No. 2.5e-05;
Matches 208; Conservative 127; Mismatches 348; Indels 320; Gaps 55;
```

```
Qy 30 IMVPOG-----IYDGETLTVSPFYVIGDPSTTVFSAGELTLKLNDSIAALPL-----79
Db 79 VMPAGTNPGNINLNDVTVN-----VNDASG---YAKG-IIIGKNSSLTANRLTVDVV 129
Qy 80 ---SCFG-NLLGSFTVLGRGHSITFENIRTSTNGAALSNSA--ADGLFTIEGFKELSPFN 133
Db 130 GQTSAGINLIGDYTHADLG--TGSTIKSNDGIIIGHSTLTATQFTIE-----N 178
Qy 134 CNSL-LAVLPAATT--NKGSTPTTST-----PSNGTI--YSKTDLLLNNEKF 178
Db 179 SNGIGLTINDYGTSDLGSGSKITDGTGVYIGLNGNNAAGAAARFTATDLTI-----232
Qy 179 SPYSLNVSDGGGAIDAKSLTVQGISKLCVFOENTAQDGGACQVTFVSFAMANEAPIAFV 238
Db 233 -----DVQYSGAMGINVQKNSVVDLGTNTSTIKTNGDNHAGLWSFCQVSANALTVDV 283
Qy 239 ----ANVAGVGGGIAAVQDQG---QGVSSSTSTEDPVVFSFRTAVFEDGNVAVGGG 290
Db 284 TGAANGVGVRCGTTIGADSHISSAQGGGLVTSSTDAINFSGTAAQR-----NS 334
Qy 291 IYSGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYDGGGAIKFCNKAQAGS 350
Db 335 IFSGSGYASAGATATAVIN-----MQNTDITVDRNLSLALGLWALSOG 377
Qy 351 NNSG-SVSPDGEVGVFFSSNVAAGKGAIA-----KKLSVANGCPVQFLRNIAND 400
Db 378 RITGSLAITG-----AAGARG-IYAMTNSQIDLTSLVIDMSTPDQWAIATQHD 426
Qy 401 GG--ATYLGESGELSADYDGIIFDGNLKR-----TAKENAADVNG-----VTVSS 445
Db 427 DGYAASRINASGEMLIN---GSVLKGGILNDMHPGVSVMTGSSLSLSDNVNGGKLDVAMNN 483
Qy 446 QAIMSGSGGKITTLR-----AKAGHOILFNDPIEMANGNQAQSSKLLKIND 493
Db 484 SVMNVTSNLSLDTLALSHSTVDFPASHGSTAGTFTTLN--VENLSGNSSTFIMRADVVGEGN 541
Qy 494 GEGYTGDIVFANGSSSTLYQWVTIEQ-----GRIVLR-----EKAKLSVNSLSOTGGS 540
Db 542 GVNNRGDLINI SGSSAGNHVLAI RNQGSBATTGNEVLTVVKTDDGAASFSSAQVBLGGY 601
Qy 541 LY--MEAGSTWDFV-----TPQP-----POQPPAANQLITLSNLHLSLANNV 585
Db 602 LYDVKNGTNWELYSAGTVPEPTNPPEPTPAQPP-----I 638
Qy 586 TNP-PTNPPAQDSHPAVIGSTTA---GSVTISGPIF-----FEDLDOTAYDRYD 630
Db 639 VNPDPTEPAPTKEP-----TTTADAGNVLAVYLLNVYENRTLMQRMGDLRNQSKDGN 694
Qy 631 WL-----GSNQKINVLKLOLQGTGK-PPANAPSDLTGLN 661
Db 695 WLRSYGGSLDSPASGKLSGDFMGYSIGQFGGDKRLSDVMPLVYGLVIDSTHASPDSGGD 754
Qy 662 EMPKYGQGSWKLAWDPNTANNPPT---LKAWTYTKTGYNPGPERSVLPVNSLWGSILD 718
Db 755 GTARSDTMGNIA-----SYMAQNGFYDLVLIKASRQKNSFH-----VLD 793
Qy 719 IRSAHSAIQASVDGRSYC-----RGLWVSGVSNFFYHRRDALGQGYRYISGG 765
Db 794 SQNNGVNANGTANGMSISLEAGQRFNLSPTGYFYIETPQTQLTYSHQENAM---KASNG 850
```


QY 766 YSLGANSYFGSMFGLAETVEFGRSKDYVCRSNHACIGSVYLSQ--QALCG--SYLP 821
DB 851 LNIHLNHY--BSLGRA-SMILG-----YDITAGNSQL--NYYVKTGAIRFSGDTEYLL 900
QY 822 GDAPIRASYGF-GN-----OHMKTSYTFABESVVRWDN 853
DB 901 NDS--REKYSFGKNGWNGVGSQAQYNK-QHTFYLEADYTQGN 940

RESULT 33

US-11-067-260-24
; Sequence 24, Application US/11067260
; Publication No. US20060051840A1
; GENERAL INFORMATION:
; APPLICANT: Arico, Maria
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISSERIAL PROTEINS
; FILE REFERENCE: 22302099501
; CURRENT APPLICATION NUMBER: US/11/067,260
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: 10/220,480
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: PCT/IB01/00420
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1392
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: deletion G983-961c
US-11-067-260-24

Query Match 3.6%; Score 190; DB 11; Length 1392;
Best Local Similarity 21.0%; Pred. No. 4.3e-05;
Matches 211; Conservative 120; Mismatches 378; Indels 296; Gaps 52;
QY 60 PSAGELTKNLNDNSIAALPLSCFNLGSLFTVLGRGHSLTF---ENIRTSNGAAL---- 112
DB 268 YSGGDKT---DEGIRLMQSDYGNL--SYHIRKNMLFIPSTGNDAAQAPNTYALLPFY 321
QY 113 SNSAADGLFTIEG-----FKELSF-----SNCSLLAVLPAATNKGSTPT 154
DB 322 EKDAQKGIITVAGVDRSGEKFREMYGPGTEPLEYGNHCGITAMWCLSAPEASVRFT 381
QY 155 TTS-----TPSNGTIYSKTDLLILNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVP 208
DB 382 RTNPIQIAGTSFSAPIVTGTAAALLL--QKYPWMSN-----DNLRRT 420
QY 209 QENTAQAQDGGACQVVTSP-----SAMANEAPIAFVANVAGVRGGIAAQQDQGVSV 260
DB 421 LLITTAQ-DIGAVGVDSKFGWGLLDGAKMNGPAGPPGDFDTADTKG-----465
QY 261 SSTSTEDPVFSFRNTAVEPQGNVARVCGGIYSVGNVAF-----LNNCKTLFLNNVASPV 315
DB 466 ----TSDIAYSF-RNDISGTGLIKKGGSQLHKNNTYTKTIEBGSGLVLYGNKSDM 520
QY 316 YIAAKQPT--SQOASNTSNNGYDGAIFCKNGAAGSNNS-----GSVSFDGEGVVFFSSN 369
DB 521 RVEYKGLIYNGAASGGSLN--SDGIVYLAQDQSGANETHVHIGSLQDQKGLIY-----574
QY 370 VAAGK-----GGAIYARK-----LSVANCQ--PVQFLRNANDG 401
DB 575 TRLGKLLKVDGTAIIGKLYMSARGKAGYLNSTGRRVPFLSAKIGQDYFFFTNIETDG 634
QY 402 GAIYLGESGELSADYGD-----IFDGNLKRATK-----ENRADVNGVTVSQA 447
DB 635 GLLASLDSVEKTAGSE-GDTLSYVRRGNRAARTASAAHSAHPAGLKHAVEQGSNLENLM 693
QY 448 ISMG---SGGKITTLRAKAGHQLLPNDPIEMANGNNQPAQSKLLKINDGEGYTGDI VFA 504
DB 694 VELDASESATPEIVETAA-----DRTDMPGIRPYGATTPRAAAVQHANAADGVRIFN 747

QY 505 NGSSTLYQNTVIE---QGRIVLREKAKLSVNSLSQSGGLYMBR-----GSTWD----- 550
DB 748 SLAATVYADSTAAHADMQGR-----RLKAVSDGLDHNHGTGLRVIAQTQDQGGTWEQGGVE 802
QY 551 -----FVTQPQPQPPAANQLITLSNLHLSSLALLANNAVTPPTNPPAQDSHPA 600
DB 803 GKMRGSTQTVGIAAKTGENTAA-----ATLGMGRSTWSENSA--NAKT-----DSLSL 849
QY 601 VIG-STTAGSVTISGPFIFFEDLDDTAYDRYD-----WLGSNQ-----KINVLKQLQGTKPP 650
DB 850 FAGIRHDAGDIGYLKGLF-----SYGRYKNSISRSTGADEHAEGSVNGTLMQLGALGG 902
QY 651 ANAP-----SDLTGN-----EMPKYGYQGSWKLAWDPTNANNPPTLKAATWTGTGYNP 699
DB 903 VNVPPAATGDLTVEGGLRYDILKQDAFAEKG-S-ALGWSGNSLREG-----946
QY 700 GPERVASIVPNSLWMSILD--IRSAHSAIQASVDGRSVCRLWLVSQVSNFFYHDDRDLAQ 757
DB 947 ---TLVGLAGLKQSPLSDKAVLPATAGVERDLNGRDYT-----VTGGFTGATAATGK 996
QY 758 -----GYRYISGYSILGANSYFGSMFGLAETVEFGRSKDYVCRSNHACIGSVYLS 810
DB 997 TGAENMPHTRLVAG---LGADVDFGNGWGLARYSYAG-SKQY-----GNHSGRVGVGY-- 1046
QY 811 TQOALCGSYLPGDA-----FIRASYGFGNQ-----HMKTSYTFABESDVVRWDN 853
DB 1047 --REFEGGGTGSATNDDVDKAAATVAIAAAYNNQBEINGFKAGETIYDIDEGTIT-KK 1103
QY 854 NCLAGEI-----CAGLPIVITPSKLYLNELRPPVQAEFSYADHE 892
DB 1104 DATAADVEADDFKGLGLKVKVTVNLTKTVNENKQNVDAKVAARSE 1148

RESULT 34

US-11-067-260-22
; Sequence 22, Application US/11067260
; Publication No. US20060051840A1
; GENERAL INFORMATION:
; APPLICANT: Arico, Maria
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISSERIAL PROTEINS
; FILE REFERENCE: 22302099501
; CURRENT APPLICATION NUMBER: US/11/067,260
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: 10/220,480
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: PCT/IB01/00420
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: deletion G983-961
US-11-067-260-22

Query Match 3.6%; Score 190; DB 11; Length 1447;
Best Local Similarity 21.0%; Pred. No. 4.5e-05;
Matches 211; Conservative 120; Mismatches 378; Indels 296; Gaps 52;
QY 60 PSAGELTKNLNDNSIAALPLSCFNLGSLFTVLGRGHSLTF---ENIRTSNGAAL---- 112
DB 268 YSGGDKT---DEGIRLMQSDYGNL--SYHIRKNMLFIPSTGNDAAQAPNTYALLPFY 321
QY 113 SNSAADGLFTIEG-----FKELSF-----SNCSLLAVLPAATNKGSTPT 154
DB 322 EKDAQKGIITVAGVDRSGEKFREMYGPGTEPLEYGNHCGITAMWCLSAPEASVRFT 381
QY 155 TTS-----TPSNGTIYSKTDLLILNNEKFSYSLVSGDGAIDAKSLTVQGISKLCVP 208
DB 382 RTNPIQIAGTSFSAPIVTGTAAALLL--QKYPWMSN-----DNLRRT 420

QY 209 QENTAAQDGCACQVTSF-----SAMANEAPAFVANVAGVGGGIAAQQDQGGVSV 260
DB 421 LTTAAQ-DIGAVGVDSKFGWLLDAGKAWGSPFPFGDTADTKG----- 465
QY 261 SSTSTEDPVVFSRNTAVEFGNVARVGGIYSYGNVAF-----LNNKTKTLFNNVSPV 315
DB 466 ----TSDIAYSP-RNDISGTGGLIKKGSQQLHGNNTYTGKTIIBGGSVLVYGNKSDM 520
QY 316 YIAAQKPT--SGOANTSNNYDGGAI PCNKAQAQGSNNS-----GSVSPDGGVGFPSN 369
DB 521 RVETKALYNGAAGCSLN--SDGIVYLADTQSGANETVHI KGSLOLDGKGTLY---- 574
QY 370 VAAK-----CGAIYAKK-----LSVANGC-PVQFLNFIANDG 401
DB 575 TRGLKLLKVDGTAIIGKLYMSARGKAGYLNGSTRVRPFLSAAKIQDYSFPTNIEDG 634
QY 402 GAIYLGESGELSADYGD---IIPDGNLKTAK-----ENAAVNVGVTVSSQA 447
DB 635 GLLASLDSVEKTAGSE-GDTLSYVVRGNAARTASAAHSAPAGLKHAVEQGGSNLENLM 693
QY 448 ISNG-----SGKITTILRAKAGHQLFNDPIEMANGNQPAQSSKLLKINDGEGTGDVFA 504
DB 694 VELDASESSATPETVETAAA-----DRTDMPGIRPYGATFRAAAAVOHANAADGVRIEN 747
QY 505 NGSSSTLYQNVTIE---QGRIVLREKAKLSVNSLSQSGSLYMEA-----GSTWD----- 550
DB 748 SLAATVYADSTAAHADMOGR-----RLKAVSDGLDHNGTGLRVIAQTQDQGGTWEQGV 802
QY 551 -----FVTPQPPQPPAAQNLITLSNLHLSSLLANNNAVTPNPPAODSHPA 600
DB 803 GMRGSGTQVGTAAKTGENTTAA-----ATLCKGRSTWSENSA--NAKT-----DSLSL 849
QY 601 VIG-STTAGSVTIGSPIPEDLDTDAYDRYD-----WLGSNQ-----KINVKLQLGTQKP 650
DB 850 FAGIRHDAGDIGYKGLF-----SYGRYKNSISRSTGADEHAGSVNGTLMQLGALGG 902
QY 651 ANAP-----SDLTGDN-----EMPKYGVQGSWKLAWDPTNANGPYTLKATWTKGYNP 699
DB 903 VNVPPAATGDLTVEGGLRYDLKQDAFAEKG--ALGWSGNSLTEG----- 946
QY 700 GPERVASLPNSLWGSILD--IRSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDALGQ 757
DB 947 ---TLVGLAGLKLQPLSDKAVLFATAGVERDLNGRDYT-----VTGGFTGATATGK 996
QY 758 -----GYRISGGYSLGANSYFGSSMFLAFTEVFGRSKDYVVCRSNHHACIGSYLS 810
DB 997 TGARNMPHTLVAG---LGADVEFGNGWGLARYSYAG--SKQY---GNHSGRVGVGY-- 1046
QY 811 TQALCGSYLFGDA-----FIRASYFGNQ-----HMKTSYTFABESDVWRDN 853
DB 1047 --RFLGGGGTGSATNDDVKAATVAIAAAYNNGQBEINGFKAGETIYDDEGTTT-KK 1103
QY 854 NCLAGEI-----GAGLPVITPSKLYNLRLPPVQAEFSYADHE 892
DB 1104 DATAADVEADDFKGLGLKRVVTLTKTVNENKONVDKAVKAESE 1148

RESULT 35

US-11-052-554A-283
; Sequence 283, Application US/11052554A
; Publication No. US20050286866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 283
; LENGTH: 1461

; TYPE: PR1

; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-283

Query Match 3.6%; Score 189.5; DB 11; Length 1461;

Best Local Similarity 20.2%; Pred. No. 5e-05;

Matches 240; Conservative 161; Mismatches 415; Indels 373; Gaps 62;

QY 20 SLNGGGYA-AEIMVPGIYDGETL-----TVSPFYTVIGDPS-----GTTVP 60

DB 203 TFGSGQAGATIQKDS--NGNTIASTQVDDNNGHWSVSLPTQSAEHTWVSVQIVGSTIT 260

QY 61 SAGELTKLNDLSIAALPLSCFG-----NLLGSFTVLGRGHSITPENIRTS 106

DB 261 DAGSITL-TIDNSQASVQVATTAGDNIINASQAAAGFTLSGTSSHLAQOQTELT-----VT 314

QY 107 TNGAALSNS-AADGLFTIEGPKELSFSCNCSLLAVLPAATTNKGSQTPPTTSTPSSNGTIY 165

DB 315 LNKGYTTSVVGANGAWSVQ-----VPTADAAQALGEGNQAVLVSGKDATGNTV 361

QY 166 SKTDLLLNNKESFYSNLVSGD---GGAIDAKSLTVQIGISKLCVPQENTAOADG----- 217

DB 362 TGAQLLTVDTPPTLAINITIAQDNIISAAEHNVALLVSGTSNAEAGQTVTLTVNGKSHA 421

QY 218 -----GACQV---VTSFSMA--NEAPIAFVANVAGVGGGIAAQQDQGGVSSSSTED 267

DB 422 TVGSDGTQWQTLTPATEVQALAEQNAVNASVSDRAG-----NTTSHSANFTVDTSA 472

QY 268 PVVSFSRNTAVEFDGNVARGVGGIYSYGNVAFLLNKGKTLFLNNVASPVYIAAKQPTSGQA 327

DB 473 PVVSV--NTVAGDD-----ILNNAEQ-----AVAQIISGVQ 501

QY 328 SNTSNNGYDGGAIFCCKNGAQ-----AGSNNSGSVSPD-----GEGVVPFSSNVAAG 373

DB 502 SGAS-----PGDVTYVKLGTHVLGTGIVLADGWNVALDPVATRTLDRGANTIFVTVTDAAG 557

QY 374 KGAAYAKKLSVANGCPQVFLNIANDG-----GAIYLGESGE-----LSL 414

DB 558 NTGAA--SRAITLVGVSPPLITINTVSGDDIISGAEGAPLTLTGSTQQAETGQTVTVTLAG 616

QY 415 SADYGDIIIDGNLKR-----AKENAAD--VNGVTSSQAI SMGSGGKITTLRAKAGHQLIF 469

DB 617 QSFTTTVQADGWSLTVPAAMGNLDPGAVAITASVTDLSGNTGNTSRITVDSQAPALS 676

QY 470 NDPIEMAN-----GNNOP-----AQSSKLLKIN--DGEQYTGDI----- 501

DB 677 IDPLTADNIIINAESQDLPIGTGTTDAQPGQTVTVTLNGQTYQGVQPDGTMVSVTPAAN 736

QY 502 --VFANGSSTLYQNVTIEQ-----RIVL-----REKAKLSVN 532

DB 737 VGLADGNATVTVASVNDVAGNPSVSRVALVDATPPVVTINPVATDINVTPEHAQAI 796

QY 533 SILSQTGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSSLLANNNAVTPNPTNP 592

DB 797 SGTVTGA---QAG---DIVT-----VTLLNNVDYTTVVDGSGNWSLGVPAV 836

QY 593 P---AQDSHPAVI-----GSTTAGSVTI-----SGPIFFEDLDDTAYDRYDMLGNSQKIN 639

DB 837 VSGLDGSPVSVSVTDKAGNTGQSLSITVTVTAAPLI--GINSIAGD--DVINASEKGA 892

QY 640 VLKQLQGTTPPANAPSDLTLG--NEMPKYQGSWKL-----AWDPNTANNOPYTLKATWTK 694

DB 893 DLQITGTSQDPVNTAITVTLNGQNTYTTTDSAGNWSVTVPAASAVTALGOANYTVTAATS 952

QY 695 TGNVGPFRVASLVFNSLWGSILD-----IRSAHSAIQASVDGRSYCRGLWVSGVSNF 747

DB 953 DIGNSATASHNVLDVSDALPGVTINPVATDIIINAAEAGVAQTISGQ-----VTGA--- 1002

QY 748 FYHDDALGQGYRYSIGGYSILGANSYFGSSMFLAFTEVFGRSKDYVVCRSNHHACIGSV 807

Matches 170; Conservative 110; Mismatches 293; Indels 219; Gaps 38;	
QY 23	GGGYAAEIMVPGIYDGETLTVSPYTVIGDPSGTTVSAGELTLKLNDSIAALPLSCF 82
DB 284	CGMYADSI-----TLIANEKGVGNAGTL-----EAAKQLIVTSS 319
QY 83	GNLIGS----FTVLGRGHSHTPENIRTSNGAA---LNSA---ADGLFTIEGPKELSPS 132
DB 320	GRIENSGRIATTADGTETASPTVLSIETTEKGAAGTIFISNGGRIESKGLLVIETGEDISLR 379
QY 133	NCNLSLAVLPAATTNKGSTPTTT--STPSNGTIYSKTDLLLNNEKFPSPYSLNVLSDGGA 191
DB 380	N-----GAVVQNGSRPAITVLNAGHLVIESKTN---VNNAKGS--ANLSA--GGR 424
QY 192	IDAKSLTVQGISKLCVFQENTAAQDGCQVVTSPSAMANEAPIAFVANAVAGVRGGGIA 251
DB 425	TTINDATIQAQSVSVYSTKGTDELGENTRIIAENVTLSNGS-----IGSAAVIEAKDTAH 480
QY 252	VODGQ-----QGVSSSTSTEDPVVSFSRNTAVEPDGNVAVRGGGIYSYGNVAFLNNGKT 305
DB 481	IESGKPLSLETSTVASNIRLNGNINIKGGKQALLADDNITAKTNTNTPGNL--YVHTGKD 539
QY 306	LFLN--NVASPVVIAAKPTSGOASNTSNNGDGAIFCKNGAQAG-----SNNSG 354
DB 540	LNINVDKDISAHSIHLKSDNAHICTSKLTASKDM---GVEAGLLNVNTNLRNTNSG 595
QY 355	SVSPDGEVYVFFSSNVAQKGG--AIYAKKLSVANCGPVQFLR--NIANDGAIYVLGSEGL 412
DB 596	NL-----HIQAAGNIQLRNTKLNAAKALETTALQCNIVSD--GLHAVSADGHV 642
QY 413	SLSADYGDILPDGNLKRITAKENADYNG-----VTVSSQAISM 450
DB 643	SLAN--GNADFTGHNTLTAK---ADVNAGSVGKRLKADNTNITSSSGDITLVAGNIGQL 698
QY 451	GGG-----GKITTIRAKAGHOILFNDPIEMANG-----NNQPAQSS----- 486
DB 699	GDGQRNSNGKHISIKNNGENADLKNLVHAKSGALNIHSDRALSIENTKLESTHNTLH 758
QY 487	----KLLKINDGEGY-----TGDIVF-----ANG-----SSTLYQNVTI 516
DB 759	NAQHERVTLNQVDAYAHRHLSITGSQIQWQNDKLP SANKLVANGVLALNARYSQIADNTTL 818
QY 517	EQRIVLREKAKL-----SVN-----SLSQTCGSLYMEAGSTWDFVTPPOPPQ 559
DB 819	RAGAINLTAGTALVRKGNINMWSTVTKTLEDNAELKPLAGRLNIEAGS--GTLTIEP--- 873
QY 560	PPAANOLITLNLHLSSLSSLLANNAVTPNPTNPPAODSH-----PAVIGST----- 605
DB 874	---ANRISAHTDISITGGKLLLSAKGNGAGAPSAQVSSLEAKGNIRLVITGETDLRSGKI 930
QY 606	TAGS-----VTISGPIPFDDDDTAYDRYDWLGS---NQKINVKLQIQLGTKPPANAPSDL- 657
DB 931	TAGKLVVATTKKLNLEAVNNSFSNVPPTQKAAELNQSKLEQQLAQLKXSSPKSLKI 990
QY 658	-TLGNEPKYGY 668
DB 991	PTLQZERDRDLAF 1002
RESULT 38	
US-11-067-260-20	
; Sequence 20, Application US/11067260	
; Publication No. US20060051840A1	
; GENERAL INFORMATION:	
; APPLICANT: Arico, Maria	
; TITLE OF INVENTION: HETEROLOGOUS EXPRESSION OF NEISSERIAL PROTEINS	
; FILE REFERENCE: 223002099501	
; CURRENT APPLICATION NUMBER: US/11/067,260	
; CURRENT FILING DATE: 2005-02-25	
; PRIOR APPLICATION NUMBER: 10/220,480	
; PRIOR FILING DATE: 2002-08-28	
; PRIOR APPLICATION NUMBER: PCT/IB01/00420	
; PRIOR FILING DATE: 2001-02-28	

Query Match	
3.6%; Score 188.5; DB 11; Length 1312;	
Beat Local Similarity 20.1%; Pred. No. Se-05;	
Matches 225; Conservative 140; Mismatches 421; Indels 331; Gaps 56;	
QY 60	PSAGSLTLKLNDSIAALPLSCFNGNLGSLFTVLGRGHSITF---ENIRTSNNGAAL----- 112
DB 268	YSGDKT---DEGIRLMQOQSDYGNL--SYHIRKNQMLFIFSTGNDAAQAPNTYALLPPY 321
QY 113	SNSAADGLFTIEG-----FKELSF-----SNCSLLAVLPAATTNKGSTPTT 154
DB 322	EKDAQKGIITVAGVDRSGEKFREMYGPGTEPLEYSGNHCGITAMWCLSAPEASVRPT 381
QY 155	TTS-----TPSNGTIYSKTDLLLNNEKFPSPYSLNVLSDGGAIDAKSLTVQGISKLCVFP 208
DB 382	RTNPQIAGTSFSAPIVTGTAAALL--QKYPMSN-----DNLRRT 420
QY 209	QENTAAQDGCQVVTSP-----SAMANEAPIAFVANAVAGVRGGGIAAVDQGGQVVS 260
DB 421	LLTTAQ--DIGAVGDSKFGWGLLDAGKAWNGPASPFPFGDTADTKG----- 465
QY 261	SSTSTEDPVVSPSRNTAVEFDGNVAVRGGGIYSYGNVAF-----LNNGKTLFLNNVASPV 315
DB 466	----TSDIAYSP-RNDISCTGGLIKKGGSQLGHGNTVTGKTIIEGGSVLVYGNKNSDM 520
QY 316	YIAAKOPT--SQOASNTSNNGDGAIFCKNGAQAGSNNS-----GSVSFDGEGVGFSSN 369
DB 521	RVETKGLIYNGAASGGSLN--SDGIVVLADTDQSGANETVHIKGSLOLDGKGLY--- 574
QY 370	VAAK-----GGAIYAKK-----LSVANGC-PVQFLRNANDG 401
DB 575	TLRGLKLVGDGTAIIGKLYMSARGKAGYLNSTGRVPFLSAAKIQDYSFTNLETDG 634
QY 402	GAIVLGESELSLSADYGD-----IIFDGNLKRITAK-----ENAAADVNGVTYSSQA 447
DB 635	GLLASLDSVEKTAGSE-GDTLSYVRRGNAARTASAAAHSAAPAGLKHAVEQCGSNLENLM 693
QY 448	ISMG---SGKITTIRAKAGHOILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFA 504
DB 694	VELDASESATPETVETAAA-----DRTDMPGIRPYGATPRAAAAVQHANAADGVRIFN 747
QY 505	NGSSTLYQNVTI-----QGRIVLREKAKLSVNSLSQTCGSLYMEA-----GSTWD----- 550
DB 748	SLAATVIADSTAHAADMOGR-----RLKAVSDGLDHNGTGLRVIAQTQDGGTWEQGVTE 802
QY 551	-----FVTPPOPPQPPAANQLITLNLHLSSLSSLLANNAVTPNPTNPPAODSHPA 600
DB 803	GMRGSGTQTVGTAKTGENTTAA-----ATLCMGHSTWSENSA--NAKT-----DSISL 849
QY 601	VIG-STTAGSVTISGPIPFDDDDTAYDRYD-----WLGSNQ-----KINVKLQIQLGTKPP 650
DB 850	PAGIRHDAGDIGYLKGLF-----SYGRYKNSISRSTGADEHAEGSVNGTLMQLGALGG 902
QY 651	ANAP-----SDLTIGN-----EMPKYGGQGWKLANDPNTANNGPYTLKATWTKTCYNP 699
DB 903	VNVPPAATGDLTVEGGLRYDLLKQDAFAEKG--ALGWSGNSLSTEG----- 946
QY 700	GPERVASLPNSLWGSILD--IRSAHSAIQASVDGRSYCKRGLWVGSVSNFFYHDDRDLAQ 757
DB 947	---TLVGLAGLKLSQPLSDKAVLFATAGVERDLNGRDYT-----VTGFTCATATGK 996
QY 758	-----GYRISGGYSLGANSYFGSSMFLAFTEVFGRSKDYVCRSNHHACIGSVYLS 810
DB 997	TGARNPHTELVLVAG---LGADVEFGNGWGLARYSVAG--SKQY---GNHSGRVG----- 1043

Qy 811 TQALGSLYFGDAPIRASYGFQNHMKTSYTPAESDVRDNNCLAGEIGAGLPVITP 870
 Db 1044 -----VGYRFLGSGGG-----VAADIGAGLADALTA 1071
 Qy 871 SKLYNELRPVQABFSAHESFTEGDQA-RAPKSGHLNLNLSVPVGVKFDRCSSSTHPN 929
 Db 1072 PLDHKDKGLSLTLDSQVRKEKLAQAQGAETKYGNGDSLNTG---KLKDKVUS----- 1123
 Qy 930 KYSFMAAYICDA-----YRTISGPTETLLSHQETMTTDAFLARHGTVVVR---G 975
 Db 1124 RPDFIRQIEVDQLITLESGEFQVVKQSHSALTAFQTEIQDSEHSGK--MVAKRQFRIG 1181
 Qy 976 SMYASLTNIEVYCHGRYEYDASRGYGLSAGSRVP 1012
 Db 1182 DIAGHTSFDKLPEGRGATYRGTAFGSD-DAGGKLTY 1217
 RESULT 39
 US-11-052-554A-6
 ; Sequence 6, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR FILING DATE: 2005-02-07
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 6
 ; LENGTH: 949
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli 0157:H7
 US-11-052-554A-6

Query Match 3.6%; Score 187.5; DB 11; Length 949;
 Best Local Similarity 20.7%; Pred. No. 3.7e-05;
 Matches 203; Conservative 120; Mismatches 350; Indels 307; Gaps 53;
 Qy 27 AAELMVPQG--IYDGETLTVSFPYTVIGDPSGTTVPFSGAGLTLKLNLSIAALPLSCFN 84
 Db 52 AADKVVQAGETVNDG-TLTNHDNQIVFGTANGMTISTGLELGPDSSENTGGQWION--GG 108
 Qy 85 LLGSFTVLGRGHSITPENIRTSNGAALSAAADGLFTIBGKLSFNSCNSLLAVLPAA 144
 Db 109 IAGNTTVTNTRGVVLE-----GGTASDVTVIRG-----GGQSLNGLAVNTTL----- 151
 Qy 145 TTNGSGOTPTTSTPSNGTIYSKTDLLLNNEKSPYSNLVSG-----DGGAI 192
 Db 152 -NNKGEQW-VHEGVATGTIINRQD-----YQSVKSGGLATGTIINTGAEGPD 198
 Qy 193 DAKSLT---VQIGSKLCVQENTAQADGACQVVTFSAMANEAPAFVANVAGVGGGI 249
 Db 199 SDNSYTCQKQVGTAEITINKGRQ-----IILFSGLDAT-LIYAGDQSVHGRAL 249
 Qy 250 -AAVDQGGQGVSSSTSTEDPVPFSRNTAVFEDG-NVARVGGGIYSYGNVAFNLNGKTLF 307
 Db 250 NTTLNGGYQYVHRD-----GLALNTVINEGGWQVVKAGG---AAGNTTINQNGE--- 295
 Qy 308 LNNVASPVYIAAKPTSQASNTSNYCDGAIKCKGAQA-GSNNSGSVFD---GEGV 363
 Db 296 -----LRVHAGGEATVATQN---TGGALVTSTAATVIGTNRLGNFTVENGKADGV 342
 Qy 364 VFFSNNVAAGKGGAIYAKLSVANCGPVQVFLRNANDGGAIIYLGESGEL-----SL 414
 Db 343 VLES-----GG-----RLDVLSSHSAQ--NTLVDDGGTLAVSAGGKATSVTITSGAL 388

Qy 415 SADYGDIIIFDGNLKRTAKENADVNGVTVSSQAISMGSCKIT-----T 458
 Db 389 IADSG-----ATVEGTMASGKPSIDGTSQASGLLENCGSFVNAGQAGNTTVGHRGT 443
 Qy 459 LRAKAGHQIILFNDPIEMANGNQPQSSKLLKINDGEGYTGDIIVFANGSSSTLYQNVTIEQ 518
 Db 444 LTLAAGGSL-----SGRTQLSKGASMV-LNGDVVSTGDIV--NAGEIRFPDNTTPN 491
 Qy 519 ---GRIVLREKA-----KLSVNSLSOTGSLYMEA---GSTWDFVTPVTPPPQPPAANQLI 567
 Db 492 AALSRAVAKNSPVTPHKLTTLNLTGCGGTINMRVLDGNS-----ASDQLV 538
 Qy 568 TLSNLHLSLSILLANNAVTPNPPPAQDHPHVIAGTSTAGSVTISGPIPEDLDDTAYD 627
 Db 539 INGGQATGKTWL-----APT-----VGNSTLGVAT----- 564
 Qy 628 RYDWLGSNQKINVLKLGTKPPANAPASDLTLGNEMPKYGVQGSKWLAWDPNTANNPPT 687
 Db 565 -----TGQIRVVDQAQNGATTE-----EGAPALS-RPLQAGAFNYT 599
 Qy 688 LKATWTKTGYNPGPVRVASLVPNSLMGSIID-----IRSAHSAIQASVDGRSYCRGLW 740
 Db 600 LNRDSDWYLRGENAVRAEVP--LYTSMLTQAMDYDRILAGSRSHQTVGNNGENSVRLS 657
 Qy 741 VSGVSNFPYHDDRALGCGYRYISGGYSLGANSYFGSSMFG-----LATTEVFGSKDY 793
 Db 658 IQG--GHLGHDN-----NGGIARGATPE--SSGSYGVRLLEGDLRTEVAGMSLT 704
 Qy 794 VVCRSHHACI-----GSVYLSLQALCGSLYFGDAFTRASYGP-----GNQH 836
 Db 705 GVTGAAGHSSVDVKDDGSRAGTV--RDDAGSLGGYL---NLVTSGLWADIVAQOTRH 759
 Qy 837 -MKTSTYTPAESDVRDNNCLAGEIGAGLPVITPSSKLYNELRPFPVQABFSAHESFT 895
 Db 760 SMKAS---SDNDFRAGWGMWLSLETGLPFSITDNL-----LEPQLQ-----YT 802
 Qy 896 BEG-----DQARAFKSGH 908
 Db 803 WQGLSLDDGQDNAGYVKFGH 822
 RESULT 40
 US-11-052-554A-284
 ; Sequence 284, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR FILING DATE: 2005-02-07
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 284
 ; LENGTH: 1343
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli 0157:H7
 US-11-052-554A-284
 Query Match 3.5%; Score 186.5; DB 11; Length 1343;
 Best Local Similarity 19.5%; Pred. No. 7.2e-05;
 Matches 174; Conservative 123; Mismatches 310; Indels 285; Gaps 43;
 Qy 15 AYSCCSLNGGGYAAEIMVPGQIYDGETLTVSFPYTVIGDPSGTTVPFSGAGLTLKLNLSI 74
 Db 494 SPDTVGINVGPNFGSGIV---NVNGATLNSTGYGFIGGNASGKGIIVNISTDSLWNLKTS 550
 Qy 75 AALPLSCFGLNLSGFTVVLGRGH-SLTIPENIRTSNGAALSAAADGLFTIEGKELSFNS 133

Db 551 TNAQLQVG-----VLGTGELNITGGIVARDTQIALNDKSKDVRVDG----- 595

QY 134 CNSLLAV--LPAATTNKGSTPTTTST-----PSNGTI-----YSKTDL 170

Db 596 QNSLLETFMVYGTSGTGLTLTNSGTLNVEGGEVVLGVFEPVAVGTLNIGAAGHGEAAADA 655

QY 171 -LLLNNKPSFYSN-----LVSGD-----GGAI 192

Db 656 GPITNATKVFEGSGEGVFVFNHTNNSDAGYQVDMILITGDDKDKVIHDAGHTVFNAGNTY 715

QY 193 DAKSLTVQGISKLCVFOENTAOADG---GACQVTFPSAMANEAPIAFVANVAG----- 243

Db 716 SKTLVNDGLLIA-----SHTADGVTGMSSEVTI---ASPGLDILASTNSAGDYTLT 767

QY 244 --VRGGIAAVQDQGGVSSSTSTEDPVVVSFRNTAVERFDG----- 282

Db 768 NALKGDGLMRVQ-----LSSDKMFGFTHATGTEPAGVAQLKMDSTFLERDNTAAL 818

QY 283 -----NVARVGGGIYSGNVAFNNGKTLFLNNVASPV-----YIAAKOPTSGQAS 328

Db 819 THAMLOSDIENTSVNVGQSIGGLA--MNGGTLIFDTPAATLABGYISVDTLVVGASD 877

QY 329 NT--SNNY--GDGGAIF--CKNGAQAGSNN--GSVSFDGE--GVVFFSSNVAAGKG 375

Db 878 YTWKGRNYQVNGRVDVLIIGVPKPNWDPMANPLTTLNLEHDDNHVGVQLVRAQTVIGSG 937

QY 376 GAIYAKKLSVANGCPVQFLNRIAND---GGAIYLGESGELSLSADYG-----DII 422

Db 938 GSL-----TLRDLOGDEVEADKTLHIAQNGTVVAEGDYGFRLLTAPGDGLY 983

QY 423 FDGNLKRKAKENADVNTVSSQAI SMGSGKITTILRAKAGHQ--ILFND--PIEMANG 478

Db 984 VNYGLK-----ALNIHGQKLTAEHGGAYGATADMSAKIGGEDLAINTVRQVLSNG 1037

QY 479 NNQPAOSSKLLKI-----NDGEGVTGD-----IVPANGSSTLYQNVITIEQGRIVLR 524

Db 1038 QND-YQGATYVQMTLRTDADGALGNTRLEINISNAALVDLNGSTQVTFETFGQGSTVLF 1096

QY 525 EKAKLSVN-----SLSQTGGSLYMEA-GSTWDFVTPQPPQPPAANQLITLS 570

Db 1097 KEGSLTVNKGGISQBELTGGGNLVNVTGTLAVEGLNARYNALTSPFNAEVSLDNTQGLG 1156

QY 571 NLHLSLSSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTIISGPIFFELDDTAYDRYD 630

Db 1157 RGNIANDGLLTLKNVTGELRN---SISGKGIVSATARTDVELDG-----DNSR 1201

QY 631 WLGSNQKINVLQLGKTPPANAPSDLTLGNEMPKYGYQGSWKLAWDPNTANNPVTL-- 688

Db 1202 FVG---QFNI-----DTGSALSV-NEQKNLG-----DASVINNGLLITIST 1237

QY 689 KATWTKTYNPGPERVASLVPSNLWGS-----ILDIRSAHSAIQASVD 731

Db 1238 ERSWAMT-----HSISGSGDLTKLGTGILTLNNDSSAYQGTTD 1275

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:11:34 ; Search time 190 Seconds
(without alignments)
2340.270 Million cell updates/sec

Title: US-10-701-844-2
Perfect score: 5267
Sequence: 1 MQTSHKFLSLAYSCS.....VEYRDASRGYGLSAGSRVRF 1012

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 21.*
1: Geneseqp1380s.*
2: Geneseqp1390s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5267	100.0	1012	2 AAY16735	Aay16735 C. tracho
2	5131.5	97.4	1013	5 ABG91021	Abg91021 Chlamydia
3	5131.5	97.4	1013	7 ADD43798	Add43798 Chlamydia
4	5131.5	97.4	1013	7 ADD43722	Add43722 Chlamydia
5	5131.5	97.4	1013	7 ADD43718	Add43718 Chlamydia
6	5131.5	97.4	1013	7 ADD43708	Add43708 Chlamydia
7	5131.5	97.4	1013	7 ADD43714	Add43714 Chlamydia
8	5131.5	97.4	1013	7 ADD43786	Add43786 Chlamydia
9	5131.5	97.4	1013	7 ADD43788	Add43788 Chlamydia
10	5131.5	97.4	1013	7 ADD43710	Add43710 Chlamydia
11	5131.5	97.4	1013	7 ADD43706	Add43706 Chlamydia
12	5131.5	97.4	1013	9 ADW29027	Adw29027 C. trachom
13	5131.5	97.4	1013	9 AEA19078	Aea19078 Chlamydia
14	5131.5	97.4	1013	9 AEA19002	Aea19002 Chlamydia
15	5125.5	97.3	1013	9 AEA19080	Aea19080 Chlamydia
16	5123.5	97.3	1013	9 AEA19006	Aea19006 Chlamydia
17	5123.5	97.3	1013	9 AEA19090	Aea19090 Chlamydia
18	5123.5	97.3	1013	9 AEA19010	Aea19010 Chlamydia
19	5121.5	97.2	1013	9 AEA19000	Aea19000 Chlamydia
20	5121.5	97.2	1013	9 AEA18998	Aea18998 Chlamydia
21	5112.5	97.0	1013	9 AEA19014	Aea19014 Chlamydia
22	5106.5	97.0	1013	2 AAY16737	Aay16737 C. tracho
23	5106.5	97.0	1013	2 AAY16738	Aay16738 C. tracho
24	5090	96.6	1006	4 ABG83207	Abg83207 Protein e

25	5090	96.6	1006	5 ABB94178	Abb94178 Chlamydia
26	5084	96.5	982	3 AAB13633	Aab13633 C. tracho
27	5084	96.5	982	4 AAG83201	Aag83201 Protein e
28	5084	96.5	982	5 ABB94172	Abb94172 Chlamydia
29	5083	96.5	1006	3 AAB13639	Aab13639 C. tracho
30	3336.5	63.3	670	7 ADD42756	Add42756 Chlamydia
31	2593	49.2	524	2 AAY37238	Aay37238 Chlamydia
32	2547	48.4	505	2 AAY16739	Aay16739 C. tracho
33	2350	44.6	458	2 AAY16751	Aay16751 Chlamydia
34	1735	32.9	325	2 AAY16752	Aay16752 Chlamydia
35	1365.5	25.9	973	5 ABB90527	Abb90527 Chlamydia
36	1359.5	25.8	973	3 AAY96274	Aay96274 Chlamydia
37	1165	22.1	918	3 AAY69369	Aay69369 Amino aci
38	1143	21.7	918	2 AAW88422	Aaw88422 Chlamydia
39	1128	21.4	928	3 AAY94327	Aay94327 Chlamydia
40	1122	21.3	926	5 ABP56019	Abp56019 Chlamydia
41	1122	21.3	926	5 ABB98228	Abb98228 Chlamydia
42	1122	21.3	926	6 ABU66284	Abu66284 C. psitt
43	1121	21.3	928	2 AAW88421	Aaw88421 Chlamydia
44	1101	20.9	928	2 AAW88418	Aaw88418 Chlamydia
45	1091	20.7	928	5 ABB90573	Abb90573 Chlamydia

ALIGNMENTS

RESULT 1
AAY16735
ID AAY16735 standard; protein; 1012 AA.
XX AC AAY16735;
XX 21-JUL-1999 (first entry)
XX DE C. trachomatis LGV L2 HMW protein.
XX KW Chlamydia; high molecular weight protein; HMW protein; urethritis;
KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
KW salpingitis; tubal occlusion; infertility; cervical cancer;
KW arteriosclerosis; atherosclerosis.
XX OS Chlamydia trachomatis.
XX PN W09917741-Al.
XX PD 15-APR-1999.
XX PF 01-OCT-1998; 98WO-US020737.
XX PR 02-OCT-1997; 97US-00942596.
XX PA (ANTE-) ANTEX BIOLOGICS INC.
XX PI Jackson JW, Pace JL;
XX DR WPI; 1999-287659/24.
XX DR N-PSDB; AAX60539.
XX PT New Chlamydia protein useful for treating conjunctivitis, urethritis and cervical cancer.
XX PS Claim 4; Fig 3; 141pp; English.
XX CC The invention relates to an isolated Chlamydia species high molecular weight (HMW) protein having an apparent mol. wt. of 105-115 kD as determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can be used for preventing, treating or ameliorating a disorder related to Chlamydia e.g. bacterial infection, conjunctivitis, urethritis, lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis, pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical cancer, infertility, arteriosclerosis and atherosclerosis. The products can also be used for detection and diagnosis. The present sequence

CC	represents a C. trachomatis HMW protein																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											</
----	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Db	961	TDAFHARHGVVVRGSMYASLTSNIEVYGHGRYEDASRGYGLSAGSRVRF	1012							
RESULT 2										
ABG91021										
ID	ABG91021	standard; protein; 1013 AA.								
XX	AC	ABG91021;								
XX	DT	29-NOV-2002 (first entry)								
XX	XX	Chlamydia trachomatis outer membrane protein G protein.								
DE	XX	Gram-negative bacterial bleb; PorB; outer membrane protein;								
KW	XX	Chlamydia trachomatis infection; Chlamydia pneumoniae infection;								
KW	XX	protective antigen; antibacterial; vaccine; gene; ds.								
OS	XX	Chlamydia trachomatis.								
XX	XX	WO200262380-A2.								
XX	XX	15-AUG-2002.								
XX	XX	08-FEB-2002; 2002WO-EP001356.								
XX	XX	08-FEB-2001; 2001GB-00003169.								
XX	XX	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.								
XX	XX	Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;								
XX	XX	WPI; 2002-657510/70.								
XX	XX	N-PSDB; ABS67342.								
XX	XX	Novel gram-negative bacterial bleb presenting on its surface PorB outer membrane protein from Chlamydia trachomatis or protective antigen from Chlamydia pneumoniae, useful for preventing Chlamydia infection.								
XX	XX	Disclosure; Page 10; 75pp; English.								
XX	XX	The present invention relates to a new gram-negative bacterial bleb presenting on its surface the PorB outer membrane protein from Chlamydia trachomatis, or a protective antigen from C. pneumoniae. The invention is useful for preventing C. trachomatis or C. pneumoniae infection in a host. The present nucleic acid sequence represents a Chlamydia trachomatis gene of the invention								
XX	XX	SQ Sequence 1013 AA;								
Query Match 97.4%; Score 5131.5; DB 5; Length 1013;										
Best Local Similarity 97.0%; Pred. No. 0;										
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;										
Qy	1	MQTSFHKFLSMILAYSCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP	60							
Db	1	MQTSFHKFLSMILAYSCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVP	60							
Qy	61	SAGELTLKNDLNSIAALPLSCFNLGSLTFVLGRGHSITPENIRTSNGAALSNSAADGL	120							
Db	61	SAGELTLKNDLNSIAALPLSCFNLGSLTFVLGRGHSITPENIRTSNGAALSNSAADGL	120							
Qy	121	FTIEGFKELSPSNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLLNNEKFSF	180							
Db	121	FTIEGFKELSPSNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLLNNEKFSF	180							
Qy	181	YSNLVSGDGGGDAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN	240							
Db	181	YSNLVSGDGGGDAIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVAN	240							
Qy	241	VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAFI	300							
Db	241	VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAFI	300							

QY 301 NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNYYGDCGAI FCKNGAQ-AGSNNSGSVSPD 359
DB 301 NNGKTLFLNNVSPVYIAAEQPTNGQASNTSDNYGDCGAI FCKNGAQAGSNNSGSVSPD 360
QY 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLRNANDGGA IYLGESGELSADYG 419
DB 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLGNANDGGA IYLGESGELSADYG 420
QY 420 DIIFDGNLKR TAKENADVNTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 479
DB 421 DIIFDGNLKR TAKENADVNTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 480
QY 480 NOPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQSG 539
DB 481 NOPAQSSSEPLKINDGEGYTGDI VFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQSG 540
QY 540 SLYMEAGSTWDFVTPOPPQPPAANQILITLSNLHLSSLLANNVTPPTPPAQDSHP 599
DB 541 SLYMEAGSTLDFVTPOPPQPPAANQILITLSNLHLSSLLANNVTPPTPPAQDSHP 600
QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKQINVLKQLGTPPPANAPSDLT 659
DB 601 AIIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKQIDVLKQLGTPPSANAPSDLT 660
QY 660 GNEMPKYGYQGSWKLAWDPNTANGPYTLKATWTKGYNPGPERVASLVPNSLWGSILDI 719
DB 661 GNEMPKYGYQGSWKLAWDPNTANGPYTLKATWTKGYNPGPERVASLVPNSLWGSILDI 720
QY 720 RASHSIAQASVDGRSVCRLGVSGVSNFFPHDRDALGQGYRYISGGYSLGANSYFGSSMF 779
DB 721 RASHSIAQASVDGRSVCRLGVSGVSNFFPHDRDALGQGYRYISGGYSLGANSYFGSSMF 780
QY 780 GLAFTEVFGRSKDYVVVCRSNHACIGSVYLSLSTQALCGSYLFGDAFIRASYGFGNQHMT 839
DB 781 GLAFTEVFGRSKDYVVVCRSNHACIGSVYLSLSTQALCGSYLFGDAFIRASYGFGNQHMT 840
QY 840 SYTFABESDVRMNNCLAGEIGAGLPVITPSKLYLNLRLPFPVQAFPSYADHESFTEGD 899
DB 841 SYTFABESDVRMNNCLAGEIGAGLPVITPSKLYLNLRLPFPVQAFPSYADHESFTEGD 900
QY 900 QARAPKSGHLNLNVSPVGVKFRDCSSHPNKYSFMAAYICDAVRTISGTTLLSHQETW 959
DB 901 QARAFPSGHLNLNVSPVGVKFRDCSSHPNKYSFMAAYICDAVRTISGTTLLSHQETW 960
QY 960 TTDAPFLARHGVVVRGSMVASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1012
DB 961 TTDAPFLARHGVVVRGSMVASLTSNIEVYGHGREYRDASRGYGLSAGSRVRF 1013

RESULT 3

ADD43798
ID ADD43798 standard; protein; 1013 AA.

AC ADD43798;

XX 15-JAN-2004 (first entry)

DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 93.

XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.

OS Chlamydia trachomatis.

XX W02003049762-A2.

XX 19-JUN-2003.

XX 12-DEC-2002; 2002WO-IB005761.

XX 12-DEC-2001; 2001GB-00029732.

PR 06-AUG-2002; 2002GB-00018233.

14-AUG-2002; 2002GB-00018924.

XX (CHIR-) CHIRON SPA.

XX Grandi G, Ratti G;

XX WPI; 2003-532882/50.

XX N-PSDB; ADD43799.

PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.

PS Claim 6; SEQ ID NO 93; 164pp; English.

XX The invention relates to a novel immunogenic composition comprising a
XX protein or nucleic acid, and an adjuvant, where the protein or nucleic
XX acid comprises any of 131 fully defined amino acid or nucleotide
XX sequences given in the specification. The protein and/or nucleic acid of
XX identity to it, or their fragments. The protein and/or nucleic acid of
XX the immunogenic composition is useful in the manufacture of a medicament
XX for the treatment or prevention of infection due to Chlamydia
XX trachomatis. The infection is treated or prevented by the medicament
XX eliciting an immune response which is specific to a C. trachomatis
XX elementary body, or for neutralising C. trachomatis elementary bodies,
XX hence the immunogenic composition can be used in creating a vaccine. The
XX immunogenic compositions can also be used for the diagnosis of C.
XX trachomatis infection. The nucleic acids of the immunogenic compositions
XX can be used to treat disorders by gene therapy. The immunogenic
XX compositions have antibacterial activity. This sequence represents one of
XX the 131 C. trachomatis proteins with immunogenic properties of the
XX invention.

SQ Sequence 1013 AA;

Query Match 97.4%; Score 5131.5; DB 7; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 993; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 MOTSFHKPFLSLMLAYSCSLNGGVAARIWVQGIYDGETLTVSPYTVIGDPSGTTVF 60
DB 1 MOTSFHKPFLSLMLAYSCSLNGGVAARIWVQGIYDGETLTVSPYTVIGDPSGTTVF 60
QY 61 SAGELTKNLNLSIALPLSCFNLIGSFTVLGRGHS�TFENIRTSNGAALSNSAAGL 120
DB 61 SAGELTKNLNLSIALPLSCFNLIGSFTVLGRGHS�TFENIRTSNGAALSNSAAGL 120
QY 121 FTIEGPKLSFNSCNLSLLAVLPAATNKGSOPTTTSTPSNGTIYKTDLLLNNEKPSF 180
DB 121 FTIEGPKLSFNSCNLSLLAVLPAATNKGSOPTTTSTPSNGTIYKTDLLLNNEKPSF 180
QY 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAAQDGGACQVVTFSAMANEAPIAFVAN 240
DB 181 YSNLVSQDGAIDAKSLTVQGISKLCVFOENTAAQDGGACQVVTFSAMANEAPIAFVAN 240
QY 241 VAGVRGGIAAVQDGGVSSSTSTEDPVVPSFRNTAVEFDGNVAVRGVGGIYSGNVAF 300
DB 241 VAGVRGGIAAVQDGGVSSSTSTEDPVVPSFRNTAVEFDGNVAVRGVGGIYSGNVAF 300
QY 301 NNGKTLFLNNVSPVYIAAKQPTSGQASNTSNYYGDCGAI FCKNGAQ-AGSNNSGSVSPD 359
DB 301 NNGKTLFLNNVSPVYIAAEQPTNGQASNTSDNYGDCGAI FCKNGAQAGSNNSGSVSPD 360
QY 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLRNANDGGA IYLGESGELSADYG 419
DB 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLGNANDGGA IYLGESGELSADYG 420
QY 420 DIIFDGNLKR TAKENADVNTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 479
DB 421 DIIFDGNLKR TAKENADVNTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 480
QY 480 NOPAQSSKLLKINDGEGYTGDI VFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQSG 539

Db 721 RSAHSRAIQASVDGRSYCRGLVWSGVNFFYHDDALGQGYRYSIGSYLSGANSYFGSSMF 780
QY 780 GLAFTVEFGRSKDYVVCRRNHACIGSVYLSVSTQOALCGSYLFGDAFIRASVYFGNQHMKT 839
Db 781 GLAFTVEFGRSKDYVVCRRNHACIGSVYLSVSTQOALCGSYLFGDAFIRASVYFGNQHMKT 840
QY 840 SYTFAESDVRWDDNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFSYADHESFTEGD 899
Db 841 SYTFAESDVRWDDNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFSYADHESFTEGD 900
QY 900 QARAFKSGHLLMLSVPGVKFDRCSSTHENKYSFMAAYICDAYRTISGTETLLSHQETW 959
Db 901 QARAFKSGHLLMLSVPGVKFDRCSSTHENKYSFMAAYICDAYRTISGTETLLSHQETW 960
QY 960 TTDAPHLARHGVIIVRGSVYASLTNSIEVYGHGRYERDTSRGYGLSAGSKVRP 1012
Db 961 TTDAPHLARHGVIIVRGSVYASLTNSIEVYGHGRYERDTSRGYGLSAGSKVRP 1013

RESULT 5
ADD43718
ID ADD43718 standard; protein; 1013 AA.
XX AC
XX ADD43718;
XX DT 15-JAN-2004 (first entry)
XX Chlamydia trachomatis immunogenic protein, SEQ ID No 13.
DE Chlamydia trachomatis immunogenic protein, SEQ ID No 13.
XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW gene therapy; antibacterial.
KW Chlamydia trachomatis.
XX OS
XX WO2003049762-A2.
XX PN
XX 19-JUN-2003.
XX PF 12-DEC-2002; 2002WO-IB005761.
XX PR 12-DEC-2001; 2001GB-00029732.
XX PR 06-AUG-2002; 2002GB-00018233.
XX PR 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
XX PA
XX Grandi G, Ratti G;
XX PI
XX WPI; 2003-532882/50.
XX DR N-PSDB; ADD43719.
XX DR

XX New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
XX
XX Claim 6; SEQ ID NO 13; 164pp; English.
XX
XX The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of

CC the 131 C. trachomatis proteins with immunogenic properties of the
CC invention.
XX
SQ Sequence 1013 AA;
Query Match 97.4%; Score 5131.5; DB 7; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;
QY 1 MQTSFHKFFLMLAYSCSLNGGGAIAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF 60
Db 1 MQTSFHKFFLMLAYSCSLNGGGAIAEIMVPOGIYDGETLTVSPFYTVIGDPSGTTVF 60
QY 61 SAGELTLKXLDNSIAALPLSCFNLGSLTFLGRHSITFENIRFTSTNGAALSNAADGL 120
Db 61 SAGELTLKXLDNSIAALPLSCFNLGSLTFLGRHSITFENIRFTSTNGAALSNAADGL 120
QY 121 FTIEGFKLSFNCNLSLAVLPAATTNNGSQTPTTTPSNGTIYKTDLLLNNEKFSF 180
Db 121 FTIEGFKLSFNCNLSLAVLPAATTNNGSQTPTTTPSNGTIYKTDLLLNNEKFSF 180
QY 181 YSNLVSGDGGAIIDAKSLTVQGISKLCVFQENTAQADGACQVTSFSAMANEAPAFVAN 240
Db 181 YSNLVSGDGGAIIDAKSLTVQGISKLCVFQENTAQADGACQVTSFSAMANEAPAFVAN 240
QY 241 VAGVRGGGIAAVQDQCGQVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
Db 241 VAGVRGGGIAAVQDQCGQVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
QY 301 NNGKTLFLNNVASPVYIAAKQPTSCQASNTSNYDGGAI FCNKAQ-AGSNNSGSVSD 359
Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSNYDGGAI FCNKAQAAGSNNSGSVSD 360
QY 360 GEGVVFSSNVAAGKGAIYAKKLSVANGCPVQFLNINANDGAIYLGSGELSLADYVG 419
Db 361 GEGVVFSSNVAAGKGAIYAKKLSVANGCPVQFLNINANDGAIYLGSGELSLADYVG 420
QY 420 DIIFDGNLKRITAKENAADVNGVTVSSQAI SMGSGGKI TTRAKAGHQLIFNDPIEMANG 479
Db 421 DIIFDGNLKRITAKENAADVNGVTVSSQAI SMGSGGKI TTRAKAGHQLIFNDPIEMANG 480
QY 480 NQPAQSSKLLKINDGEYTGDI VFPANGSTLYQNTTIEQRIVLREKAKLSVNSLSQTGG 539
Db 481 NQPAQSSKLLKINDGEYTGDI VFPANGSTLYQNTTIEQRIVLREKAKLSVNSLSQTGG 540
QY 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPPTPPAODSHP 599
Db 541 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNAVTPPTPPAODSHP 600
QY 600 AVIGSTTAGSVTISGPIPPEDLDYDRYDMLGNSQKINVLKQLGKTKPPANAPSDLT 659
Db 601 AVIGSTTAGSVTISGPIPPEDLDYDRYDMLGNSQKIDVLKQLGKTKPPANAPSDLT 660
QY 660 GNEMPKYGGQSKWLAWDPNTANNNGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDI 719
Db 661 GNEMPKYGGQSKWLAWDPNTANNNGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDI 720
QY 720 RSAHSAIQASVDGRSYCRGLVWSGVNFFYHDDALGQGYRYSIGSYLSGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSYCRGLVWSGVNFFYHDDALGQGYRYSIGSYLSGANSYFGSSMF 780
QY 780 GLAFTVEFGRSKDYVVCRRNHACIGSVYLSVSTQOALCGSYLFGDAFIRASVYFGNQHMKT 839
Db 781 GLAFTVEFGRSKDYVVCRRNHACIGSVYLSVSTQOALCGSYLFGDAFIRASVYFGNQHMKT 840
QY 840 SYTFAESDVRWDDNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFSYADHESFTEGD 899
Db 841 SYTFAESDVRWDDNCLAGEICAGLPVITTPSKLYLNLRLPFPVQAEFSYADHESFTEGD 900
QY 900 QARAFKSGHLLMLSVPGVKFDRCSSTHENKYSFMAAYICDAYRTISGTETLLSHQETW 959
Db 901 QARAFKSGHLLMLSVPGVKFDRCSSTHENKYSFMAAYICDAYRTISGTETLLSHQETW 960

QY	960	TTDAPHLARHGIVVGRGSMYASLTSLNIEVYGHGRYERDASRCGYLSAGSRVRF	1012
Db	961	TTDAPHLARHGIVVGRGSMYASLTSLNIEVYGHGRYERDASRCGYLSAGSKVRP	1013
RESULT 6			
ADD43708			
ID	ADD43708	standard; protein; 1013 AA.	
XX	AC	ADD43708;	
XX	15-JAN-2004	(first entry)	
DE	Chlamydia trachomatis immunogenic protein, SEQ ID No 3.		
XX	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;		
KW	gene therapy; antibacterial.		
XX	Chlamydia trachomatis.		
OS	WO2003049762-A2.		
PN	19-JUN-2003.		
XX	12-DEC-2002; 2002WO-IB005761.		
XX	12-DEC-2001; 2001GB-00029732.		
PR	06-AUG-2002; 2002GB-00018233.		
PR	14-AUG-2002; 2002GB-00018924.		
XX	(CHIR-) CHIRON SPA.		
PA	Grandi G, Ratti G;		
PI	MPi; 2003-532882/50.		
DR	N-PSDB; ADD43709.		
DR	New immunogenic composition having a protein or encoding nucleic acid,		
PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis		
PT	infection.		
XX	Claim 6; SEQ ID NO 3; 164pp; English.		
XX	The invention relates to a novel immunogenic composition comprising a		
CC	protein or nucleic acid, and an adjuvant, where the protein or nucleic		
CC	acid comprises any of 131 fully defined amino acid or nucleotide		
CC	sequences given in the specification, or has 50% or greater sequence		
CC	identity to it, or their fragments. The protein and/or nucleic acid of		
CC	the immunogenic composition is useful in the manufacture of a medicament		
CC	for the treatment or prevention of infection due to Chlamydia		
CC	trachomatis. The infection is treated or prevented by the medicament		
CC	eliciting an immune response which is specific to a C. trachomatis		
CC	elementary body, or for neutralising C. trachomatis elementary bodies.		
CC	hence the immunogenic composition can be used in creating a vaccine. The		
CC	immunogenic compositions can also be used for the diagnosis of C.		
CC	trachomatis infection. The nucleic acids of the immunogenic compositions		
CC	can be used to treat disorders by gene therapy. The immunogenic		
CC	compositions have antibacterial activity. This sequence represents one of		
CC	the 131 C. trachomatis proteins with immunogenic properties of the		
CC	invention.		
XX	Sequence 1013 AA;		
QY	960	TTDAPHLARHGIVVGRGSMYASLTSLNIEVYGHGRYERDASRCGYLSAGSRVRF	1012
Db	961	TTDAPHLARHGIVVGRGSMYASLTSLNIEVYGHGRYERDASRCGYLSAGSKVRP	1013
Query Match			
Best Local Similarity 97.4%; Score 5131.5; DB 7; Length 1013;			
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;			
QY	1	MQT9FHKFPLSMILAYSCCSLGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVF	60
Db	1	MQT9FHKFPLSMILAYSCCSLGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPSGTTVF	60
QY	61	SAGELTKNLNDNSTAALPLSCFNLGSGFTVLGRGHSITFENIRSTNGAALSASDGL	120
Db	61	SAGELTKNLNDNSTAALPLSCFNLGSGFTVLGRGHSITFENIRSTNGAALSASDGL	120
QY	121	FTIEGFKELSPSCNCSLLAVLPAATTNKGSOTPTTSTPSNGTIYSKTDLLLNNEKFSF	180
Db	121	FTIEGFKELSPSCNCSLLAVLPAATTNKGSOTPTTSTPSNGTIYSKTDLLLNNEKFSF	180
QY	181	YSNLVSGDGGGAIIDAKSLTVQGISKLCVPQENTAQDGGACQVVTFSFSAANEAPAFVAN	240
Db	181	YSNLVSGDGGGAIIDAKSLTVQGISKLCVPQENTAQDGGACQVVTFSFSAANEAPAFVAN	240
QY	241	VAGVRGGGIAAVQDQGGVSSSTSTEDPVVPSFSTAVEPDGNVAVCGGIYSYGNVAF	300
Db	241	VAGVRGGGIAAVQDQGGVSSSTSTEDPVVPSFSTAVEPDGNVAVCGGIYSYGNVAF	300
QY	301	NGKTLFLNNVASPVYIAAQPTSCQASNTSNNTDGGAIKCKGAG-AGSNNGSVSPD	359
Db	301	NGKTLFLNNVASPVYIAAQPTSCQASNTSNNTDGGAIKCKGAG-AGSNNGSVSPD	359
QY	360	GEVVPFSSNVAAGKGAIVAKKLSVANGCQVQPLRNIANDGGAIYLGESGELSADYG	419
Db	361	GEVVPFSSNVAAGKGAIVAKKLSVANGCQVQPLRNIANDGGAIYLGESGELSADYG	420
QY	420	DIIFDGNLKRITAKENAADVNGVTYSSQAIKSGSGGKITTIRAKACHQIILFNDPIEMANG	479
Db	421	DIIFDGNLKRITAKENAADVNGVTYSSQAIKSGSGGKITTIRAKACHQIILFNDPIEMANG	480
QY	480	NQPAQSSKLLKINDGEGTGDIVPANGSSTLYQNVTTIQQRIVLREKAKLSVNSLSQTGG	539
Db	481	NQPAQSSKLLKINDGEGTGDIVPANGSSTLYQNVTTIQQRIVLREKAKLSVNSLSQTGG	540
QY	540	SLYNEAGSTWDFVTPQPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPQDSHP	599
Db	541	SLYNEAGSTWDFVTPQPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPQDSHP	600
QY	600	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSGNQKINVLKQLGTQKPPANAPSDTL	659
Db	601	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDNLGSGNQKIDVLKQLGTQKPPANAPSDTL	660
QY	660	GNEMPKYGYQGSWKLAWDPNTANNPPTLKATWTKGYNPQPERVASLVPNSLWGSILDI	719
Db	661	GNEMPKYGYQGSWKLAWDPNTANNPPTLKATWTKGYNPQPERVASLVPNSLWGSILDI	720
QY	720	RSASAIQASVDGRSICRGVSGVSNFFYHDDRDLQGYRISGGYSLGANSYFGSSMF	779
Db	721	RSASAIQASVDGRSICRGVSGVSNFFYHDDRDLQGYRISGGYSLGANSYFGSSMF	780
QY	780	GLAFTEVFGRSKDYVCRSNHACIGSVYLSLTOQALCGSVLFGDAFTRASVFGNQHMKT	839
Db	781	GLAFTEVFGRSKDYVCRSNHACIGSVYLSLTOQALCGSVLFGDAFTRASVFGNQHMKT	840
QY	840	SYTPAESDVRWNNCLAGIBIGAGLPITVITPSKLYLNLRLPFFVQAEFSYADHESFTSEGD	899
Db	841	SYTPAESDVRWNNCLAGIBIGAGLPITVITPSKLYLNLRLPFFVQAEFSYADHESFTSEGD	900
QY	900	QARAFKSGHLNLTGVPGVGFDRCSSTHPNKYSFMAAYICDAYNTISGTETTLSSHQETW	959
Db	901	QARAFKSGHLNLTGVPGVGFDRCSSTHPNKYSFMAAYICDAYNTISGTETTLSSHQETW	960
QY	960	TTDAPHLARHGIVVGRGSMYASLTSLNIEVYGHGRYERDASRCGYLSAGSRVRF	1012
Db	961	TTDAPHLARHGIVVGRGSMYASLTSLNIEVYGHGRYERDASRCGYLSAGSKVRP	1013
RESULT 7			
ADD43714			
ID	ADD43714	standard; protein; 1013 AA.	
XX	AC	ADD43714;	
XX	15-JAN-2004	(first entry)	
DE	Chlamydia trachomatis immunogenic protein, SEQ ID No 9.		

KW	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW	gene therapy; antibacterial.
XX	
OS	Chlamydia trachomatis.
XX	
PN	WO2003049762-A2.
XX	
PD	19-JUN-2003.
XX	
PF	12-DEC-2002; 2002WO-IB005761.
XX	
PR	12-DEC-2001; 2001GB-00029732.
PR	06-AUG-2002; 2002GB-00018233.
PR	14-AUG-2002; 2002GB-00018924.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PI	Grandi G, Ratti G;
XX	
DR	WPI; 2003-532882/50.
DR	N-PSDB; ADD43715.
XX	
PT	New immunogenic composition having a protein or encoding nucleic acid,
PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT	infection.
XX	
XX	Claim 6; SEQ ID NO 9; 164pp; English.
PS	
CC	The invention relates to a novel immunogenic composition comprising a
CC	protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC	acid comprises any of 131 fully defined amino acid or nucleotide
CC	sequences given in the specification, or has 50% or greater sequence
CC	identity to it, or their fragments. The protein and/or nucleic acid of
CC	the immunogenic composition is useful in the manufacture of a medicament
CC	for the treatment or prevention of infection due to Chlamydia
CC	trachomatis. The infection is treated or prevented by the medicament
CC	eliciting an immune response which is specific to a C. trachomatis
CC	elementary body, or for neutralising C. trachomatis elementary bodies,
CC	hence the immunogenic composition can be used in creating a vaccine. The
CC	immunogenic compositions can also be used for the diagnosis of C.
CC	trachomatis infection. The nucleic acids of the immunogenic compositions
CC	can be used to treat disorders by gene therapy. The immunogenic
CC	compositions have antibacterial activity. This sequence represents one of
CC	the 131 C. trachomatis proteins with immunogenic properties of the
CC	invention.
XX	
SQ	Sequence 1013 AA;
Query Match 97.4%; Score 5131.5; DB 7; Length 1013;	
Best Local Similarity 97.0%; Pred. No. 0;	
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;	
QY	1 MQTSEHKKFLLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTTVP 60
DB	1 MQTSEHKKFLLSMILAYSCCLNGGGYAAEIMVPGIYDGETLTVSPFTYVIGDPSGTTVP 60
QY	61 SAGELTLKLNLDNSIAALPLSCFNGLLGSFTVLGRGHSLLTPENIRTSNGAALNSAADGL 120
DB	61 SAGELTLKLNLDNSIAALPLSCFNGLLGSFTVLGRGHSLLTPENIRTSNGAALSDSANSGL 120
QY	121 FTTEGFKELSFNSCNLSLAVLPAATTNKGSQTPTTTPSNGTIYSKTDILLANKEKFSF 180
DB	121 FTTEGFKELSFNSCNLSLAVLPAATTNKGSQTPTTTPSNGTIYSKTDILLANKEKFSF 180
QY	181 YSNLVSGDGGAIIDAKSLTVQGISLKVFOENTAGDGCQVTVTSFSAMANEAPAFVAN 240
DB	181 YSNLVSGDGGAIIDAKSLTVQGISLKVFOENTAGDGCQVTVTSFSAMANEAPAFIAN 240
QY	241 VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVVSFSRNTAVFEDGNVARYGGGIYSYGNVAF 300
DB	241 VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVVSFSRNTAVFEDGNVARYGGGIYSYGNVAF 300
QY	301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNTYDGGGAIFCKNGAQ-AGSNNSGSVSFD 359

DB	301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNTYDGGGAIFCKNGAAGSNNSGSVSFD 360
QY	360 GEGVVFFSSNVAAGKGGAIYAKKLSVANGCPVQFVLRNIANDGGAIYLGESGELSADYG 419
DB	361 GEGVVFFSSNVAAGKGGAIYAKKLSVANGCPVQFVLRNIANDGGAIYLGESGELSADYG 420
QY	420 DIIFDGNLKRTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLFNDPIEMANGN 479
DB	421 DIIFDGNLKRTAKENADVNGVTSSQAI SMGSGGKITTLRAKAGHQLFNDPIEMANGN 480
QY	480 NOPAQSLLKINDGEGYTGDI VFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQTGG 539
DB	481 NOPAQSSEPLKINDGEGYTGDI VFANGSSTLYQNTYIEQGRIVLREKAKLSVNSLSQTGG 540
QY	540 SLYMEAGSTWDFVTVPQPPQPPAANQLITLSNLHLSLSLLANNNAVTPPTPPAQDSDP 599
DB	541 SLYMEAGSTWDFVTVPQPPQPPAANQLITLSNLHLSLSLLANNNAVTPPTPPAQDSDP 600
QY	600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKINVLKQLGTPKPPANAPSDLT 659
DB	601 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDWLGSNQKIDVLKQLGTPKPPANAPSDLT 660
QY	660 GNEMPKYVQGSWKLAWDPNTANNGPYTLKATWTKTYNPGPERSVASLVPNSLWGSILDI 719
DB	661 GNEMPKYVQGSWKLAWDPNTANNGPYTLKATWTKTYNPGPERSVASLVPNSLWGSILDI 720
QY	720 RSAHSAIQASVDGRSVCRLGVSGVSNPPYHDDRDALGQGYRYISGGYSLGANSYFGSSMF 779
DB	721 RSAHSAIQASVDGRSVCRLGVSGVSNPPYHDDRDALGQGYRYISGGYSLGANSYFGSSMF 780
QY	780 GLAFTVEVFGRSKDYVVCRSNHHACIGSVYLSLSTQALCGSYLFGDAPFIRASYGFGNQHMT 839
DB	781 GLAFTVEVFGRSKDYVVCRSNHHACIGSVYLSLSTQALCGSYLFGDAPFIRASYGFGNQHMT 840
QY	840 SYTPAESDVRWNNCLAGEIGAGLPVITTPSKLYNELRPPVQAEFSYADHESFTEEGD 899
DB	841 SYTPAESDVRWNNCLAGEIGAGLPVITTPSKLYNELRPPVQAEFSYADHESFTEEGD 900
QY	900 QARAFKSGHLLNLVSPGVKFCRSTHBNKYSEMAAYICDAVRTISGTTETLLSHQETW 959
DB	901 QARAFRSGHMLNLSVPGVKFCRSTHBNKYSEMAAYICDAVRTISGTTETLLSHQETW 960
QY	960 TTDAFLARHGTVVVRGSMYASLTSNIEVYGHGRYEYRDASRGYGLSAGSRVRF 1012
DB	961 TTDAFLARHGTVVVRGSMYASLTSNIEVYGHGRYEYRDTSRGYGLSAGSKVRF 1013
RESULT 8	
ADD43786	
ID	ADD43786 standard; protein; 1013 AA.
XX	
AC	ADD43786;
XX	
DT	15-JAN-2004 (first entry)
XX	
DE	Chlamydia trachomatis immunogenic protein, SEQ ID No 81.
XX	
KW	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
KW	gene therapy; antibacterial.
XX	
OS	Chlamydia trachomatis.
XX	
PN	WO2003049762-A2.
XX	
PD	19-JUN-2003.
XX	
PF	12-DEC-2002; 2002WO-IB005761.
XX	
PR	12-DEC-2001; 2001GB-00029732.
PR	06-AUG-2002; 2002GB-00018233.
PR	14-AUG-2002; 2002GB-00018924.
XX	

PA	(CHIR-) CHIRON SPA.	QY	540	SLYNEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNVNTPTNPPAQDSHP	599
XX		Db	541	SLYNEAGSTLDFVTPQPPQPPAANQLITLSNLHLSLSSLLANNVNTPTNPPAQDSHP	600
XX	Grandi G, Ratti G;	QY	600	AVIGSTTAGSVTISGPIFFEDLDOTAYDRYDMLGSKNQKINVLKQLGKTPPANAPSDLTLL	659
DR	WPI; 2003-532882/50.	Db	601	AIIGSTTAGSVTISGPIFFEDLDOTAYDRYDMLGSKNQKIDVLKQLGKTPPANAPSDLTLL	660
XX	N-PSDB; ADD43787.	QY	660	GNEMPKYGYOGSKWLAWDPNTANNPVTTLKATWTKTYNPGPERSVASLVPNSLWGSILDI	719
PT	New immunogenic composition having a protein or encoding nucleic acid;	Db	661	GNEMPKYGYOGSKWLAWDPNTANNPVTTLKATWTKTYNPGPERSVASLVPNSLWGSILDI	720
PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis	QY	720	RSASHSATQASVDGSRYSYCRGLWVSGVSNFFYHRRDALCGQYRYSYSGYSLGANSYFGSSMF	779
XX	infection.	Db	721	RSASHSATQASVDGSRYSYCRGLWVSGVSNFFYHRRDALCGQYRYSYSGYSLGANSYFGSSMF	780
PS	Claim 6; SEQ ID NO 81; 164pp; English.	QY	780	GLAFTEYFGSRKDYVVCRSNHHACIGSVYLSSTQALCGSVLFGDAFIRASVYFGNQHMKT	839
XX	The invention relates to a novel immunogenic composition comprising a	Db	781	GLAFTEYFGSRKDYVVCRSNHHACIGSVYLSSTQALCGSVLFGDAFIRASVYFGNQHMKT	840
CC	protein or nucleic acid, and an adjuvant, where the protein or nucleic	QY	840	SYTPAESDVRWDMNCLAGSIFGAGLPITVITPSKLYLNLRLPFFVQAFSADHESFTTEGD	899
CC	acid comprises any of 131 fully defined amino acid or nucleotide	Db	841	SYTPAESDVRWDMNCLAGSIFGAGLPITVITPSKLYLNLRLPFFVQAFSADHESFTTEGD	900
CC	sequences given in the specification, or has 50% or greater sequence	QY	900	QARAFKSHLLNLSPVGVKFDRCSSVTHPNKYSFMAAYICDAYRTISGTETILLSHQETW	959
CC	identity to it, or their fragments. The protein and/or nucleic acid of	Db	901	QARAFKSHLLNLSPVGVKFDRCSSVTHPNKYSFMAAYICDAYRTISGTETILLSHQETW	960
CC	the immunogenic composition is useful in the manufacture of a medicament	QY	960	TTDAFLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSRVRF	1012
CC	for the treatment or prevention of infection due to Chlamydia	Db	961	TTDAFLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSRVRF	1013
CC	trachomatis. The infection is treated or prevented by the medicament	RESULT 9			
CC	eliciting an immune response which is specific to a C. trachomatis	ADD43788			
CC	elementary body, or for neutralising C. trachomatis elementary bodies,	ID	ADD43788	standard; protein; 1013 AA.	
CC	hence the immunogenic composition can be used in creating a vaccine. The	XX	AC	ADD43788;	
CC	immunogenic compositions can also be used for the diagnosis of C.	XX	DT	15-JAN-2004 (first entry)	
CC	trachomatis infection. The nucleic acids of the immunogenic compositions	XX	DE	Chlamydia trachomatis immunogenic protein, SEQ ID NO 83.	
CC	can be used to treat disorders by gene therapy. The immunogenic	XX	KW	immunogenic; infection; Chlamydia trachomatis; immune; vaccine;	
CC	compositions have antibacterial activity. This sequence represents one of	XX	OS	gene therapy; antibacterial.	
CC	the 131 C. trachomatis proteins with immunogenic properties of the	XX	PN	Chlamydia trachomatis.	
XX	invention.	XX	PD	WO2003049762-A2.	
XX	Sequence 1013 AA;	XX	PF	19-JUN-2003.	
		XX	PR	12-DEC-2002; 2002WO-IB005761.	
		XX	PR	12-DEC-2001; 2001GB-00029732.	
		XX	PR	06-AUG-2002; 2002GB-00018233.	
		XX	PR	14-AUG-2002; 2002GB-00018924.	
		XX	PA	(CHIR-) CHIRON SPA.	
		XX	PI	Grandi G, Ratti G;	
		XX	PI	WPI; 2003-532882/50.	
		XX	DR	N-PSDB; ADD43789.	
		XX	XX	New immunogenic composition having a protein or encoding nucleic acid;	
		XX	PT	useful for diagnosing, preventing and/or treating Chlamydia trachomatis	
		XX	PT	infection.	
		XX	XX	Claim 6; SEQ ID NO 83; 164pp; English.	
		XX	PS	The invention relates to a novel immunogenic composition comprising a	
		XX	CC	protein or nucleic acid, and an adjuvant, where the protein or nucleic	
		XX	CC	acid comprises any of 131 fully defined amino acid or nucleotide	
		XX	CC	sequences given in the specification, or has 50% or greater sequence	
		XX	CC	identity to it, or their fragments. The protein and/or nucleic acid of	
		XX	CC	the immunogenic composition is useful in the manufacture of a medicament	
		XX	CC	for the treatment or prevention of infection due to Chlamydia	
		XX	CC	trachomatis. The infection is treated or prevented by the medicament	
		XX	CC	eliciting an immune response which is specific to a C. trachomatis	
		XX	CC	elementary body, or for neutralising C. trachomatis elementary bodies,	
		XX	CC	hence the immunogenic composition can be used in creating a vaccine. The	
		XX	CC	immunogenic compositions can also be used for the diagnosis of C.	
		XX	CC	trachomatis infection. The nucleic acids of the immunogenic compositions	
		XX	CC	can be used to treat disorders by gene therapy. The immunogenic	
		XX	CC	compositions have antibacterial activity. This sequence represents one of	
		XX	CC	the 131 C. trachomatis proteins with immunogenic properties of the	
		XX	CC	invention.	
		XX	CC	Sequence 1013 AA;	

Query Match	97.4%;	Score 5131.5;	DB 7;	Length 1013;
Best Local Similarity	97.0%;	Pred. No. 0;		
Matches 983;	Conservative 18;	Mismatches 11;	Indels 1;	Gaps 1;
QY	1	MQTSFHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPPYTVIGDPSGTTVF	60	
Db	1	MQTSFHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPPYTVIGDPSGTTVF	60	
QY	61	SAGELTLKNLDNSIAALPLSCFNGLLGSFTVLGRHSLTPENIRKSTNGAALSNAADGL	120	
Db	61	SAGELTLKNLDNSIAALPLSCFNGLLGSFTVLGRHSLTPENIRKSTNGAALSNAADGL	120	
QY	121	FTIEGFKELSFNCNLSLAVLPAATNNKSGTPTTTSTPSNGTLYSKTDLLLLNNEKPSF	180	
Db	121	FTIEGFKELSFNCNLSLAVLPAATNNKSGTPTTTSTPSNGTLYSKTDLLLLNNEKPSF	180	
QY	181	YSNLVSGDGGAIIDAKSLITVQGISKLCVPOENTAQADGGACQVTSFSMAANEAPFAFAN	240	
Db	181	YSNLVSGDGGAIIDAKSLITVQGISKLCVPOENTAQADGGACQVTSFSMAANEAPFAFAN	240	
QY	241	VAGVGGGIAAVDQCGQGVSSSTSTEDPVVSFSRNTAVEPDGNVARVGGGIYSGNVAFL	300	
Db	241	VAGVGGGIAAVDQCGQGVSSSTSTEDPVVSFSRNTAVEPDGNVARVGGGIYSGNVAFL	300	
QY	301	NNGKTLFLNNVASPVYIAAQPTSGQASNTSNNTYDGGGAIFCKNGAQ-AGSNNSGVSFD	359	
Db	301	NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGGAIFCKNGAQAGSNNSGVSFD	360	
QY	360	GEGVVFPSSNVAAGKGAIIYAKKLSVANCQVPQFLRNANDGGGAIYLGESGELSADYG	419	
Db	361	GEGVVFPSSNVAAGKGAIIYAKKLSVANCQVPQFLRNANDGGGAIYLGESGELSADYG	420	
QY	420	DIIFDGNLKTAKENAAVNGVTYSSQAI SMGSGGKITTLRAKAGHQLFNDDPIEMANGN	479	
Db	421	DIIFDGNLKTAKENAAVNGVTYSSQAI SMGSGGKITTLRAKAGHQLFNDDPIEMANGN	480	
QY	480	NPAQSGSKLLKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG	539	
Db	481	NPAQSGSEPKINDGEGYTGDIVFANGSSTLYQNVTIEQGRIVLREKAKLSVNSLSQTGG	540	

CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This sequence represents one of
CC the 131 C. trachomatis proteins with immunogenic properties of the
CC invention.

XX
SQ Sequence 1013 AA;

Query Match		97.4%;	Score 5131.5;	DB 7;	Length 1013;
Best Local Similarity		97.0%;	Pred. No. 0;		
Matches 983;		Conservative 18;	Mismatches 11;	Indels 1;	Gaps 1;
Qy	1	MQTSFHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPGGTVF	60		
Db	1	MQTSFHKPFLSMILAYSCSLNGGGYAAEIMVPOGIYDGETLTVSPFYTVIGDPGGTVF	60		
Qy	61	SAGELTLKMLDMSIALPLSCFNLGSLTFTVGRHSLTFENIRTSNGAALSNAADGL	120		
Db	61	SAGELTLKMLDMSIALPLSCFNLGSLTFTVGRHSLTFENIRTSNGAALSNAADGL	120		
Qy	121	FTIEGPKELSPNCNSLLAVLPAATNKGSOPTTTTSPSNGTIYSKTDLLLNNEKFSF	180		
Db	121	FTIEGPKELSPNCNSLLAVLPAATNKGSOPTTTTSPSNGTIYSKTDLLLNNEKFSF	180		
Qy	181	YSLNLSGDCGADAKSLTVQGISKLCVFOENTAQAQDGGACQVVTSPSAMANEAPAFVAN	240		
Db	181	YSLNLSGDCGADAKSLTVQGISKLCVFOENTAQAQDGGACQVVTSPSAMANEAPAFVAN	240		
Qy	241	VAGVRGGGIAAVDQGGQVSSSTSTEDPVVPSRNTAVTFDGNVAVGGGIYSYGNVAF	300		
Db	241	VAGVRGGGIAAVDQGGQVSSSTSTEDPVVPSRNTAVTFDGNVAVGGGIYSYGNVAF	300		
Qy	301	NGKTLFLNNVAPVYIAAKQPTSGOASNTSNNGDGGALFCKNGAQ-AGSNNSGSVDF	359		
Db	301	NGKTLFLNNVAPVYIAAEQPTNGQAQASNTSDNYGDDGALFCKNGAQAGSNNSGSVDF	360		
Qy	360	GEGVWFFSSNVAAGKGAIAKLSVANGCPVFLRNIAANDGGAIVLGSSELSLSDY	419		
Db	361	GEGVWFFSSNVAAGKGAIAKLSVANGCPVFLRNIAANDGGAIVLGSSELSLSDY	420		
Qy	420	DIIFDGNLKRITAKENAADVNGVTVSSQAISMGGKITTIRAKAGHOILFNDPIEWANG	479		
Db	421	DIIFDGNLKRITAKENAADVNGVTVSSQAISMGGKITTIRAKAGHOILFNDPIEWANG	480		
Qy	480	NQPAQSSKLLKINDGEGYTGDIYPANGSSTLYQNVITEQRIVLRKAKLSVNSLSQTGG	539		
Db	481	NQPAQSSSEPLKINDGEGYTGDIYPANGSSTLYQNVITEQRIVLRKAKLSVNSLSQTGG	540		
Qy	540	SLYWEAGSTWDFVTPPOPPQPPAANOLITLSNLHLSLSLLANNVTPNTPNPAQDSHP	599		
Db	541	SLYWEAGSTLDFVTPPOPPQPPAANOLITLSNLHLSLSLLANNVTPNTPNPAQDSHP	600		
Qy	600	AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNINVLKQLGTPKANAPSDTL	659		
Db	601	AIIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSKNIDVLKQLGTPKANAPSDTL	660		
Qy	660	GNEMPKYQGSWKLAWDPTANNPPTLKATWTKTYNPGPERVASLVPNSLWGSLLDI	719		
Db	661	GNEMPKYQGSWKLAWDPTANNPPTLKATWTKTYNPGPERVASLVPNSLWGSLLDI	720		
Qy	720	RSASAIQASVDCGRSYCRGLVWGSVSNFFVHRRDALGQGVRYISGGYSLGANSYFGSSMF	779		
Db	721	RSASAIQASVDCGRSYCRGLVWGSVSNFFVHRRDALGQGVRYISGGYSLGANSYFGSSMF	780		

Qy	780	GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLSTQOALCGSYLFGDAFTRASYGFGNQHMKT	839
Db	781	GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLSTQOALCGSYLFGDAFTRASYGFGNQHMKT	840
Qy	840	SYTFAESDVWDNNCLAGELGAGLPVITPSPKLYLNEIRPFVQAEFSYADHESFTEBGD	899
Db	841	SYTFAESDVWDNNCLVGEIGVGLPVIPTPSKLYLNEIRPFVQAEFSYADHESFTEBGD	900
Qy	900	QARAFKSHLNLNLSVPVGVKFRDCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959
Db	901	QARAFKSHLNLNLSVPVGVKFRDCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960
Qy	960	TTDAFHLARHGCVVVRGSMYASLTSNIEVYHGRYEYRDASRGYGLSAGSRVP	1012
Db	961	TTDAFHLARHGCVVVRGSMYASLTSNIEVYHGRYEYRDTSRGYGLSAGSKVRP	1013

RESULT 10

ADD43710

ID ADD43710 standard; protein; 1013 AA.

AC ADD43710;

DT 15-JAN-2004 (first entry)

DE Chlamydia trachomatis immunogenic protein, SEQ ID No 5.

KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
gene therapy; antibacterial.

OS Chlamydia trachomatis.

PN WO2003049762-A2.

PD 19-JUN-2003.

PP 12-DEC-2002; 2002WO-IB005761.

PR 12-DEC-2001; 2001GB-00029732.

PR 06-AUG-2002; 2002GB-00018233.

PR 14-AUG-2002; 2002GB-00018924.

PA (CHIR-) CHIRON SPA.

PI Grandi G, Ratti G;

DR WP1; 2003-532882/50.

DR N-PSDB; ADD43711.

XX New immunogenic composition having a protein or encoding nucleic acid,
useful for diagnosing, preventing and/or treating Chlamydia trachomatis
infection.

XX Claim 6; SEQ ID NO 5; 164pp; English.

XX The invention relates to a novel immunogenic composition comprising a
protein or nucleic acid, and an adjuvant, where the protein or nucleic
acid comprises any of 131 fully defined amino acid or nucleotide
sequences given in the specification, or has 50% or greater sequence
identity to it, or their fragments. The protein and/or nucleic acid of
the immunogenic composition is useful in the manufacture of a medicament
for the treatment or prevention of infection due to Chlamydia
trachomatis. The infection is treated or prevented by the medicament
eliciting an immune response which is specific to a C. trachomatis
elementary body, or for neutralising C. trachomatis elementary bodies,
hence the immunogenic composition can be used in creating a vaccine. The
immunogenic compositions can also be used for the diagnosis of C.
trachomatis infection. The nucleic acids of the immunogenic compositions
can be used to treat disorders by gene therapy. The immunogenic
compositions have antibacterial activity. This sequence represents one of
the 131 C. trachomatis proteins with immunogenic properties of the
invention.

QY 121 FTIEGFKELSPNCNLSLAVLPAATNTKGSQPTTTSTPSNGTIYSKTDLLLLNNEKFSF 180
 DB 121 FTIEGFKELSPNCNLSLAVLPAATNTKGSQPTTTSTPSNGTIYSKTDLLLLNNEKFSF 180
 QY 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPIAFVAN 240
 DB 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPIAFVAN 240
 QY 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFPGDGNVAVGGGIYSYGNVAF 300
 DB 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFPGDGNVAVGGGIYSYGNVAF 300
 QY 301 NNGKTLFLNNVSPVYIAAKOPTSQASNTSNNGYDGGAI FCKNGAQ-AGSNNSSGSVSD 359
 DB 301 NNGKTLFLNNVSPVYIAAKOPTSQASNTSNNGYDGGAI FCKNGAQ-AGSNNSSGSVSD 359
 QY 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLNRIANDGGAIVLGSSELGSADYG 419
 DB 361 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLNRIANDGGAIVLGSSELGSADYG 420
 QY 420 DIIFDGNLKRKTAKENADVNGVTSSQAISMGSGGKIITLRAKAGHOILFNDPIEMANG 479
 DB 421 DIIFDGNLKRKTAKENADVNGVTSSQAISMGSGGKIITLRAKAGHOILFNDPIEMANG 480
 QY 480 NPAQSSKLLKINDGEGYTDIVFANGSSTLYONVTIEQRIIVLREKAKLSVNSLSQTTG 539
 DB 481 NPAQSSKLLKINDGEGYTDIVFANGSSTLYONVTIEQRIIVLREKAKLSVNSLSQTTG 540
 QY 540 SLVMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQSHP 599
 DB 541 SLVMEAGSTWDFVTPPOPPQPPAANQLITLSNLHLSLSSLLANNAVTPNPPAQSHP 600
 QY 600 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNQKINVLKQLGTQPPANAPSDLT 659
 DB 601 AVIGSTTAGSVTISGPIFFEDLDDTAYDRYDMLGSNQKINVLKQLGTQPPANAPSDLT 660
 QY 660 GNEMPKYQGSWKLAWDPNTANGPYTLKATWTKGYNPGERVASLVPNSLWGSLLDI 719
 DB 661 GNEMPKYQGSWKLAWDPNTANGPYTLKATWTKGYNPGERVASLVPNSLWGSLLDI 720
 QY 720 RSAHSAIOASVDCGRSYCRGLWVGSVNFYVHDDALGOGYRIYSGGYSLGANSYFGSSMF 779
 DB 721 RSAHSAIOASVDCGRSYCRGLWVGSVNFYVHDDALGOGYRIYSGGYSLGANSYFGSSMF 780
 QY 780 GLAFTEVFGSRKDYVVCRSNHHACIGSVYLSSTQALCGSVLFGDAPFRASYGFGNQHMT 839
 DB 781 GLAFTEVFGSRKDYVVCRSNHHACIGSVYLSSTQALCGSVLFGDAPFRASYGFGNQHMT 840
 QY 840 SYTPAESDVRWNNCLAGEIGAGLPVITPISKLYLNEARPFVQAEFSYADHESFTEEGD 899
 DB 841 SYTPAESDVRWNNCLAGEIGAGLPVITPISKLYLNEARPFVQAEFSYADHESFTEEGD 900
 QY 900 QARAFKSGHLLNLSVPVGVKFDRCSSHPNKSYPMAAYICDAVRTISGTETTLISHQETW 959
 DB 901 QARAFKSGHLLNLSVPVGVKFDRCSSHPNKSYPMAAYICDAVRTISGTETTLISHQETW 960
 QY 960 TTDAFHLARHGVVVRGSMYASLTNSNIEVYGHGRYEYRDAISRGYGLSAGSRVF 1012
 DB 961 TTDAFHLARHGVVVRGSMYASLTNSNIEVYGHGRYEYRDAISRGYGLSAGSRVF 1013

RESULT 12

ID ADW29027 standard; protein; 1013 AA.
 XX AC ADW29027;
 XX DT 07-APR-2005 (first entry)
 XX DE C₂trachomatis polymorphic membrane protein (PmpG) seq 21.
 XX immunogenicity; antigen; vaccine; chlamydia trachomatis infection;
 KW antibacterial; gene therapy.

XX Chlamydia trachomatis.
 OS WO2005002619-A2.
 XX 13-JAN-2005.
 XX 25-JUN-2004; 2004WO-US020491.
 XX 26-JUN-2003; 2003GB-00015020.
 XX 25-AUG-2003; 2003US-0497649P.
 XX 02-FEB-2004; 2004GB-00002236.
 XX 01-JUN-2004; 2004US-0576375P.
 XX (CHIR) CHIRON CORP.
 XX Grandi G, Finco O, Ratti G, Bonci A;
 XX WPI; 2005-075653/08.
 XX An immunogenic composition for preventing or treating Chlamydia
 trachomatis infections comprises a combination of C. trachomatis
 antigens, such as PepA, LcrB, ArtJ, DnaK, CT398, Omph-like, L7/L12, OmcA,
 or Atos.
 XX Example 1; SEQ ID NO 21; 110pp; English.
 XX This invention relates to novel immunogenic compositions that comprise a
 combination of Chlamydia trachomatis (C. trachomatis) antigens. PepA,
 Specifically, it refers to a first antigen group consisting of PepA,
 LcrB, ArtJ, DnaK and CT398 and a second antigen group consisting of PepA,
 LcrB, ArtJ, DnaK and CT398, Omph-like, L7/L12, OmcA, Atos, CT547, Enolase,
 HtrA and MrgG. The present invention describes the development of a
 vaccine and antibodies that recognize the immunogenic composition in a
 order to provide a method of neutralizing a C. trachomatis infection in a
 mammal through raising an immune response. Accordingly, the
 pharmaceutical compositions derived thereof exhibit antibacterial
 activity and the group of antigenic DNA sequences can also be used in
 gene therapy. This polypeptide sequence is a Chlamydia trachomatis
 antigen of the invention.
 XX Sequence 1013 AA;
 QY Query Match 97.4%; Score 5131.5; DB 9; Length 1013;
 DB Best Local Similarity 97.0%; Pred. No. 0;
 QY Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;
 1 MOTSFHKFPLSMILAYSCCSLNGGGYAAEIMVPOQGIYDGETLTVSPPYTVIGDPSGTTVF 60
 1 MOTSFHKFPLSMILAYSCCSLNGGGYAAEIMVPOQGIYDGETLTVSPPYTVIGDPSGTTVF 60
 61 SAGELTKNLDNSIALPLSCFNLGSGFTVLGRGHSHTFENIRTSNGAALSNSAADGL 120
 61 SAGELTKNLDNSIALPLSCFNLGSGFTVLGRGHSHTFENIRTSNGAALSNSAADGL 120
 121 FTIEGFKELSPNCNLSLAVLPAATNTKGSQPTTTSTPSNGTIYSKTDLLLLNNEKFSF 180
 121 FTIEGFKELSPNCNLSLAVLPAATNTKGSQPTTTSTPSNGTIYSKTDLLLLNNEKFSF 180
 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPIAFVAN 240
 181 YSNLVSGDGGGAIIDAKSLTVQGISKLCVFOENTAQADGGACQVVTSPSAMANEAPIAFVAN 240
 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFPGDGNVAVGGGIYSYGNVAF 300
 241 VAGVRGGGIAAVODGQGVSSSTSTEDPVVFSRNTAVFPGDGNVAVGGGIYSYGNVAF 300
 301 NNGKTLFLNNVSPVYIAAKOPTSQASNTSNNGYDGGAI FCKNGAQ-AGSNNSSGSVSD 359
 301 NNGKTLFLNNVSPVYIAAKOPTSQASNTSNNGYDGGAI FCKNGAQ-AGSNNSSGSVSD 360
 360 GEGVVFSSNVAAGKGGAIYAKKLSVANGCPVQFLNRIANDGGAIVLGSSELGSADYG 419

Db 361 GGVVFPSSNVAAGKGAIIYAKKLSVANGCPVQPLGNIANDDGAIYILGESGELSLSADYG 420
Qy 420 DIIFDGNLKRRTAKENAAADVGVTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 479
Db 421 DIIFDGNLKRRTAKENAAADVGVTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 480
Qy 480 NPAQSSKLLKINDGEGYTDIVFANGSTLYQNTIEQGRIVLREKAKLSVNSLSQGG 539
Db 481 NPAQSSKLLKINDGEGYTDIVFANGSTLYQNTIEQGRIVLREKAKLSVNSLSQGG 540
Qy 540 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599
Db 541 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTIISGPIFFEDDDTAYDRYDNLGSGNOKINVLKQLGTPPANAPSDTL 659
Db 601 AVIGSTTAGSVTIISGPIFFEDDDTAYDRYDNLGSGNOKIDVLKQLGTPPANAPSDTL 660
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNPPTLKATWTKTGYNPGERVASLVPNSLMSIIDI 719
Db 661 GNEMPKYGYQGSWKLAWDPNTANNPPTLKATWTKTGYNPGERVASLVPNSLMSIIDI 720
Qy 720 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDALGQGYRIYSGYSLGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDALGQGYRIYSGYSLGANSYFGSSMF 780
Qy 780 GLAFTVEFGRSNDYVVCRSNHHACIGSVYLSFQOALCGSYLFGDAFIRASYFGNQHMKT 839
Db 781 GLAFTVEFGRSNDYVVCRSNHHACIGSVYLSFQOALCGSYLFGDAFIRASYFGNQHMKT 840
Qy 840 SVTPAEEDVRDNNCLAGICAGLPVITPESKLYLNEILRPPVQAEPSYADHESFTREGD 899
Db 841 SVTPAEEDVRDNNCLAGICAGLPVITPESKLYLNEILRPPVQAEPSYADHESFTREGD 900
Qy 900 QARAFKSGHLMLSVFVGKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETW 959
Db 901 QARAFKSGHLMLSVFVGKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETW 960
Qy 960 TTDAFHLARHGIVVRGSMYASLTSNIYVGHGRYRYDRASRGYGLSAGSRVRF 1012
Db 961 TTDAFHLARHGIVVRGSMYASLTSNIYVGHGRYRYDRASRGYGLSAGSRVRF 1013

RESULT 13
ABAI9078
ID ABAI9078 standard; protein; 1013 AA.
XX
AC ABAI9078;
XX
DT 28-JUL-2005 (first entry)
XX
DE Chlamydia trachomatis protein - SEQ ID 81.
XX
KW chlamydia trachomatis infection; antibacterial; vaccine.
XX
OS Chlamydia trachomatis.
XX
PN US2005106162-A1.
XX
PD 19-MAY-2005.
XX
PF 16-DEC-2004; 2004US-00498327.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002EP-00182233.
PR 14-AUG-2002; 2001EP-00218924.
PR 12-DEC-2002; 2002WO-IB005761.
XX
PA (GRAN/) GRANDI G.
PA (RATTI/) RATTI G.
XX
PI Grandi G, Ratti G;
XX

DR MPI; 2005-354777/36.
DR N-FSDB; ABAI9079.
XX
PT New immunogenic composition comprising Chlamydia trachomatis protein or
PT its encoding nucleic acid, useful for immunizing against, or treating or
PT preventing chlamydial infection.
XX
PS Claim 13; SEQ ID NO 81; 90pp; English.
XX
CC The invention comprises an immunogenic composition for the treatment or
CC prevention of a Chlamydia trachomatis infection. The immunogenic
CC composition contains a Chlamydia trachomatis protein or its encoding
CC nucleic acid. The immunogenic composition of the invention is useful for
CC treating or preventing chlamydial infection. The present amino acid
CC sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was
CC obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX
SQ Sequence 1013 AA;

Query Match 97.4%; Score 5131.5; DB 9; Length 1013;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;
Qy 1 MQTSFHKPFLSMILAYSCCSLNGGGYAAIMVPOGIYDGETLTVSPFVTVIGDPSGTTVP 60
Db 1 MQTSFHKPFLSMILAYSCCSLNGGGYAAIMVPOGIYDGETLTVSPFVTVIGDPSGTTVP 60
Qy 61 SAGELTKNLNDNSTAALPLSCFGLNLSFTVLGRHSLTPENIRTSNTGAALSNAADGL 120
Db 61 SAGELTKNLNDNSTAALPLSCFGLNLSFTVLGRHSLTPENIRTSNTGAALSNAADGL 120
Qy 121 FTIEGFKELSPNCNLSLAVLPAAATNKGSTPTTSTPSNGTITYSKTDLLLANNEKPSF 180
Db 121 FTIEGFKELSPNCNLSLAVLPAAATNKGSTPTTSTPSNGTITYSKTDLLLANNEKPSF 180
Qy 181 YSNLVSGDGGAIKASLTQGISKLCYFOENTAODGACQVTVSPSANEAFIAPVAN 240
Db 181 YSNLVSGDGGAIKASLTQGISKLCYFOENTAODGACQVTVSPSANEAFIAPVAN 240
Qy 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSNTAVEPDGNVARVGGGIYSYGNVAPL 300
Db 241 VAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSNTAVEPDGNVARVGGGIYSYGNVAPL 300
Qy 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAIKPCNGAQAAGSNNSGVSFD 359
Db 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYDGGAIKPCNGAQAAGSNNSGVSFD 360
Qy 360 GEGVVPSSNVAAGKGAIIYAKKLSVANGCPVQPLGNIANDDGAIYILGESGELSLSADYG 419
Db 361 GEGVVPSSNVAAGKGAIIYAKKLSVANGCPVQPLGNIANDDGAIYILGESGELSLSADYG 420
Qy 420 DIIFDGNLKRRTAKENAAADVGVTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 479
Db 421 DIIFDGNLKRRTAKENAAADVGVTVSSQAI SMGSGGKITTLRAKAGHOILFNDPIEMANGN 480
Qy 480 NPAQSSKLLKINDGEGYTDIVFANGSTLYQNTIEQGRIVLREKAKLSVNSLSQGG 539
Db 481 NPAQSSKLLKINDGEGYTDIVFANGSTLYQNTIEQGRIVLREKAKLSVNSLSQGG 540
Qy 540 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 599
Db 541 SLYMEAGSTWDFVTPPPQPPAAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTIISGPIFFEDDDTAYDRYDNLGSGNOKINVLKQLGTPPANAPSDTL 659
Db 601 AVIGSTTAGSVTIISGPIFFEDDDTAYDRYDNLGSGNOKIDVLKQLGTPPANAPSDTL 660
Qy 660 GNEMPKYGYQGSWKLAWDPNTANNPPTLKATWTKTGYNPGERVASLVPNSLMSIIDI 719
Db 661 GNEMPKYGYQGSWKLAWDPNTANNPPTLKATWTKTGYNPGERVASLVPNSLMSIIDI 720

QY 720 RSAHSAIQASVDRGSRVCRGLWVGVSNNPFYHDDALGQGYRISGYSIGANSYFGSSMF 779
DB 721 RSAHSAIQASVDRGSRVCRGLWVGVSNNPFYHDDALGQGYRISGYSIGANSYFGSSMF 780
QY 780 GLAFTVEVFGSRKDYVVCNHNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMT 839
DB 781 GLAFTVEVFGSRKDYVVCNHNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMT 840
QY 840 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAEFSYADHESFTEGD 899
DB 841 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAEFSYADHESFTEGD 900
QY 900 QARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 959
DB 901 QARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 960
QY 960 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSKVRP 1012
DB 961 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSKVRP 1013

RESULT 14
AEA19002
ID AEA19002 standard; protein; 1013 AA.
AC AEA19002;
XX
DT 28-JUL-2005 (first entry)
XX
DE Chlamydia trachomatis protein - SEQ ID 5.
XX
KW chlamydia trachomatis infection; antibacterial; vaccine.
XX
OS Chlamydia trachomatis.
XX
PN US2005106162-A1.
XX
PD 19-MAY-2005.
XX
PF 16-DEC-2004; 2004US-00498327.
XX
PR 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002EP-00182233.
PR 14-AUG-2002; 2001EP-00218924.
PR 12-DEC-2002; 2002WO-IB005761.
XX
PA (GRAN/) GRANDI G.
PA (RATT/) RATTI G.
XX
PI Grandi G, Ratti G;
XX
XX WPI; 2005-354777/36.
DR N-PSDB; AEA19003.
XX
XX New immunogenic composition comprising Chlamydia trachomatis protein or
PT its encoding nucleic acid, useful for immunizing against, or treating or
PT preventing chlamydial infection.
XX
XX Claim 13; SEQ ID NO 5; 90pp; English.
XX
XX The invention comprises an immunogenic composition for the treatment or
CC prevention of a Chlamydia trachomatis infection. The immunogenic
CC composition contains a Chlamydia trachomatis protein or its encoding
CC nucleic acid. The immunogenic composition of the invention is useful for
CC treating or preventing chlamydial infection. The present amino acid
CC sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was
CC obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.
XX
SQ Sequence 1013 AA;

Query Match 97.4%; Score 5131.5; DB 9; Length 1013;

Best Local Similarity 97.0%; Pred. No. 0;
Matches 983; Conservative 18; Mismatches 11; Indels 1; Gaps 1;
QY 1 MQTSPHKEFLMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFVTVLGDPSGTTVF 60
DB 1 MQTSPHKEFLMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFVTVLGDPSGTTVF 60
QY 61 SAGELTLKNLONSIAALPLSCFGLNLLGSFTVLGRHSLTFENIRTSNGAALSAAADGL 120
DB 61 SAGELTLKNLONSIAALPLSCFGLNLLGSFTVLGRHSLTFENIRTSNGAALSAAADGL 120
QY 121 FTIEGFKELSFNCNLSLAVLPAATTNKSQTPPTTSTPSNGTIYSKTDLLILANKEFSF 180
DB 121 FTIEGFKELSFNCNLSLAVLPAATTNKSQTPPTTSTPSNGTIYSKTDLLILANKEFSF 180
QY 181 YSNLYSGGGGDAIDAKSLTVQGISKLCVFOENTAQADGACQVVTFSAMANEAPAFVAN 240
DB 181 YSNLYSGGGGDAIDAKSLTVQGISKLCVFOENTAQADGACQVVTFSAMANEAPAFVAN 240
QY 241 VAGVGGGGIAAVQDQGGVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
DB 241 VAGVGGGGIAAVQDQGGVSSSTSTEDPVVSFSRNTAVEFDGNVARVGGGIYSYGNVAF 300
QY 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKFCNGAQ-AGSNNSGSVSFD 359
DB 301 NNGKTLFLNNVASPVYIAAKOPTSQASNTSNNGYDGGAIKFCNGAQ-AGSNNSGSVSFD 360
QY 360 GEGVYVFFSSNVAAGKGGAIYAKKLSVANGCPVQFIGNIANDGGAIYLGSGSLSADYVG 419
DB 361 GEGVYVFFSSNVAAGKGGAIYAKKLSVANGCPVQFIGNIANDGGAIYLGSGSLSADYVG 420
QY 420 DIIIPGNLKRITAKENAADVNGVTVSSQAISSMGSGGKIITLRAKAGHQLLNDPIEMANG 479
DB 421 DIIIPGNLKRITAKENAADVNGVTVSSQAISSMGSGGKIITLRAKAGHQLLNDPIEMANG 480
QY 480 NQPAQSSKLLKINDGEGYTDIVFANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSQTCG 539
DB 481 NQPAQSSKLLKINDGEGYTDIVFANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSQTCG 540
QY 540 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTPTPPAODSHP 599
DB 541 SLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTPTPPAODSHP 600
QY 600 AVIGSTTAGSVTISGPIPFEDDDTDYDRYDMLGSKINVLKQLGKTGPPANAPSDTL 659
DB 601 AVIGSTTAGSVTISGPIPFEDDDTDYDRYDMLGSKINVLKQLGKTGPPANAPSDTL 660
QY 660 GNEMPKYGGQSWKLAWDNTANNPPTLKATWTGTGYNPGERVASLVPNSLWGSILDI 719
DB 661 GNEMPKYGGQSWKLAWDNTANNPPTLKATWTGTGYNPGERVASLVPNSLWGSILDI 720
QY 720 RSAHSAIQASVDRGSRVCRGLWVGVSNNPFYHDDALGQGYRISGYSIGANSYFGSSMF 779
DB 721 RSAHSAIQASVDRGSRVCRGLWVGVSNNPFYHDDALGQGYRISGYSIGANSYFGSSMF 780
QY 780 GLAFTVEVFGSRKDYVVCNHNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMT 839
DB 781 GLAFTVEVFGSRKDYVVCNHNHACIGSVYLSYQALCGSYLFGDAFIRASYFGNQHMT 840
QY 840 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAEFSYADHESFTEGD 899
DB 841 SYTFABESDVRDNNCLAGEIGAGLPVITPSSKLYLNELRPFVQAEFSYADHESFTEGD 900
QY 900 QARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 959
DB 901 QARAFKSGHLLNLSPVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTTLLSHQETW 960
QY 960 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSKVRP 1012
DB 961 TTDAFHLARHGVIIVRGSMYASLTNSIEVYGHGRYEYRDASRGYGLSAGSKVRP 1013

RESULT 15

CC	NOTE: The present sequence is not shown in the specification, but was									
CC	obtained from the USPTO website -									
CC	seqdata.uspto.gov/sequence.html?DocID=20050106162.									
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										
XX										

Db 481 NPAQSSBPLKINDGEGYTDIVPANGNSTLYQNVTTIQGRIVLRKAKLSVNSLSQTGG 540
Qy 540 SLYMEAGSTWDFVTPPQQPPAANQLITLSNLHLSSLLANNAVTPNPPAQDSHP 599
Db 541 SLYMEAGSTLDFVTPPQQPPAANQLITLSNLHLSSLLANNAVTPNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKINVLKQLGKTPPANAPSDTL 659
Db 601 AIIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKIDVLKQLGTPPANAPSDTL 660
Qy 660 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTGTGYNPGPERSVASLVPNSLWGSILDI 719
Db 661 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTGTGYNPGPERSVASLVPNSLWGSILDI 720
Qy 720 RSAHSAIOASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRISGGYSLGANSYFGSSMF 779
Db 721 RSAHSAIOASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRISGGYSLGANSYFGSSMF 780
Qy 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASVYFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASVYFGNQHMKT 840
Qy 840 SYTFAEEDVRWNNCLAGEICAGLPIVITPISKLYINELRPFVQAEFSDHESFTEBGD 899
Db 841 SYTFAEEDVRWNNCLAGEICAGLPIVITPISKLYINELRPFVQAEFSDHESFTEBGD 900
Qy 900 QARAFKSHLLNLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959
Db 901 QARAFKSHLLNLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 960
Qy 960 TTDAPHLARHGVIYVGRSMYASLTSNIYVGHGRYEYRDSRGYGLSAGSVRF 1012
Db 961 TTDAPHLARHGVIYVGRSMYASLTSNIYVGHGRYEYRDSRGYGLSAGSVRF 1013

RESULT 20

AEA18998
ID AEA18998 standard; protein; 1013 AA.
XX AEA18998;
AC AEA18998;
DT 28-JUL-2005 (first entry)
XX Chlamydia trachomatis protein - SEQ ID 1.
XX Chlamydia trachomatis infection; antibacterial; vaccine.
XX Chlamydia trachomatis.
OS Chlamydia trachomatis.
FN US2005106162-A1.
PD 19-MAY-2005.
PF 16-DEC-2004; 2004US-00498327.
XX 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002EP-00182233.
PR 14-AUG-2002; 2001EP-00218924.
PR 12-DEC-2002; 2002WO-IB005761.
XX (GRAN/) GRANDI G.
PA (RATT/) RATTI G.
XX Grandi G, Ratti G;
FI Grandi G, Ratti G;
XX WPI; 2005-354777/36.
DR N-P8DB; AEA18999.
XX New immunogenic composition comprising Chlamydia trachomatis protein or
PT its encoding nucleic acid, useful for immunizing against, or treating or
PT preventing chlamydial infection.
XX Claim 13; SEQ ID NO 1; 90pp; English.

XX The invention comprises an immunogenic composition for the treatment or
CC prevention of a Chlamydia trachomatis infection. The immunogenic
CC composition contains a Chlamydia trachomatis protein or its encoding
CC nucleic acid. The immunogenic composition of the invention is useful for
CC treating or preventing chlamydial infection. The present amino acid
CC sequence represents a Chlamydia trachomatis protein of the invention.
CC NOTE: The present sequence is not shown in the specification, but was
CC obtained from the USPTO website -
CC seqdata.uspto.gov/sequence.html?DocID=20050106162.

SQ Sequence 1013 AA;

Query Match 97.2%; Score 5121.5; DB 9; Length 1013;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 978; Conservative 23; Mismatches 11; Indels 1; Gaps 1;

Qy 1 MQTSFHKPFLSMILAYSCSLNGGYYAAEIMVPGIYDGETLTVSPVTVIGDPSGTTVP 60
Db 1 MQTSFHKPFLSMILAYSCSLNGGYYAAEIMVPGIYDGETLTVSPVTVIGDPSGTTVP 60
Qy 61 SAGELTLKNDNSIAALPLSCFGLGSLFTVLRGHSHTFENIRTSNGAALSAAADGL 120
Db 61 SAGELTLKNDNSIAALPLSCFGLGSLFTVLRGHSHTFENIRTSNGAALSAAADGL 120
Qy 121 FTISGPKELSPSNCNSLLAVLPAATTNKGSTPTTSTPSTNGTIYSKTDLLLNNKPSF 180
Db 121 FTISGPKELSPSNCNSLLAVLPAATTNKGSTPTTSTPSTNGTIYSKTDLLLNNKPSF 180
Qy 181 YSNLVSGDGGAIKAKSLTVQGISKLCVPOENTAQDGAAGCOVTPSPAMANEAPAVAN 240
Db 181 YSNLVSGDGGAIKAKSLTVQGISKLCVPOENTAQDGAAGCOVTPSPAMANEAPAVAN 240
Qy 241 VAGVGGGIAAVDQGGVSSSTSTEDPVVSFSNTAVFEDGNVAVGGGIYSYGNVAF 300
Db 241 VAGVGGGIAAVDQGGVSSSTSTEDPVVSFSNTAVFEDGNVAVGGGIYSYGNVAF 300
Qy 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYDGGAIKFKNGAQ-AGSNNSGVSFD 359
Db 301 NNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYDGGAIKFKNGAQ-AGSNNSGVSFD 360
Qy 360 GEGVVPFSSVAAKGGAIVAKKLSVANCGPVQPLRNANDGGAIYIGESGELSADYG 419
Db 361 GEGVVPFSSVAAKGGAIVAKKLSVANCGPVQPLRNANDGGAIYIGESGELSADYG 420
Qy 420 DIIFDGNLKRATAKENAADVNGVTVSSQAIKNGSGGKITTLRAKAGHQLFNDPIEMANG 479
Db 421 DIIFDGNLKRATAKENAADVNGVTVSSQAIKNGSGGKITTLRAKAGHQLFNDPIEMANG 480
Qy 480 NQPAQSSKLLKINDGEGYTDIVPANGNSTLYQNVTTIQGRIVLRKAKLSVNSLSQTGG 539
Db 481 NQPAQSSKLLKINDGEGYTDIVPANGNSTLYQNVTTIQGRIVLRKAKLSVNSLSQTGG 540
Qy 540 SLYMEAGSTWDFVTPPQQPPAANQLITLSNLHLSSLLANNAVTPNPPAQDSHP 599
Db 541 SLYMEAGSTLDFVTPPQQPPAANQLITLSNLHLSSLLANNAVTPNPPAQDSHP 600
Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKINVLKQLGKTPPANAPSDTL 659
Db 601 AIIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGNSQKIDVLKQLGTPPANAPSDTL 660
Qy 660 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTGTGYNPGPERSVASLVPNSLWGSILDI 719
Db 661 GNEMPKYQGSWKLAWDPNTANNPPTLKATWTGTGYNPGPERSVASLVPNSLWGSILDI 720
Qy 720 RSAHSAIOASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRISGGYSLGANSYFGSSMF 779
Db 721 RSAHSAIOASVDGRSVCRLWVSGVSNFFYHDDRDLAQGGYRISGGYSLGANSYFGSSMF 780
Qy 780 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASVYFGNQHMKT 839
Db 781 GLAFTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASVYFGNQHMKT 840

QY	840	SYTFAESDVRDNNCLAGEIGAGLPVITPSKLYLNEIRPFVQAEFSYADHESFTEGD	899
Db	841	SYTFAESDVRDNNCLGELGVGLPVIITPSKLYLNEIRPFVQAEFSYADHESFTEGD	900
QY	900	QARAFKSGHLLNLSPVGVKDFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959
Db	901	QARAFKSGHLLNLSPVGVKDFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960
QY	960	TTDAFHARHGIVVRGSMYASLTSNIEVYGHGRYEDRSRGYGLSAGSRVRF	1012
Db	961	TTDAFHARHGIVVRGSMYASLTSNIEVYGHGRYEDRSRGYGLSAGSKVRF	1013
RESULT 21			
ID	ABAI9014	standard; protein; 1013 AA.	
XX	AC	AEA19014;	
XX	28-JUL-2005	(first entry)	
DE	Chlamydia trachomatis protein - SEQ ID 17.		
XX	chlamydia trachomatis infection; antibacterial; vaccine.		
XX	Chlamydia trachomatis.		
XX	US2005106162-A1.		
XX	19-MAY-2005.		
XX	16-DEC-2004; 2004US-00498327.		
XX	12-DEC-2001; 2001GB-00029732.		
PR	06-AUG-2002; 2002EP-00182233.		
PR	14-AUG-2002; 2001EP-00218924.		
PR	12-DEC-2002; 2002WO-1B005761.		
XX	(GRAN/) GRANDI G.		
PA	(RATT/) RATTI G.		
XX	Grandi G, Ratti G;		
XX	WPI; 2005-354777/36.		
DR	N-PSDB; AEA19015.		
XX	New immunogenic composition comprising Chlamydia trachomatis protein or its encoding nucleic acid, useful for immunizing against, or treating or preventing chlamydial infection.		
XX	Claim 13; SEQ ID NO 17; 90pp; English.		
PS	The invention comprises an immunogenic composition for the treatment or prevention of a Chlamydia trachomatis infection. The immunogenic composition contains a Chlamydia trachomatis protein or its encoding nucleic acid. The immunogenic composition of the invention is useful for treating or preventing chlamydial infection. The present amino acid sequence represents a Chlamydia trachomatis protein of the invention.		
CC	NOTE: The present sequence is not shown in the specification, but was obtained from the USPTO website - seqdata.uspto.gov/sequence.html?docID=20050106162.		
XX	Seqdata.uspto.gov/sequence.html?docID=20050106162.		
XX	Sequence 1013 AA;		
Query Match 97.1%; Score 5112.5; DB 9; Length 1013;			
Best Local Similarity 96.4%; Pred. No. 0;			
Matches 977; Conservative 23; Mismatches 12; Indels 1; Gaps 1;			
QY	1	MQTSFHKFFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFYTIVGDPGGTTFV	60
Db	1	MQTSFHKFFLSMILAYSCCSLNGGGYAAEIMVPGIYDGETLTVSPFYTIVGDPGGTTFV	60
QY	61	SAGELTLKNDLSIALPLSCFGNLLGSFTVLGRGHSLTPTENIRTSNGAALSASADGL	120

Db	61	SAGELTLKNDLSIALPLSCFGNLLGSFTVLGRGHSLTPTENIRTSNGAALSASADGL	120
QY	121	FTIEGFKELSFSCNCSLLAVLPAATTNKGSTPTTSTPNSNGTIYSKTDLLLLNKEFSF	180
Db	121	FTIEGFKELSFSCNCSLLAVLPAATTNKGSTPTTSTPNSNGTIYSKTDLLLLNKEFSF	180
QY	181	YSNLVSGDGGDAIDAKSLTVQGISKLCVQFQNTAQADGGACQVTSFSAANEAPAFVAN	240
Db	181	YSNLVSGDGGDAIDAKSLTVQGISKLCVQFQNTAQADGGACQVTSFSAANEAPAFVAN	240
QY	241	VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVCGGIYSYGNVAPL	300
Db	241	VAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVCGGIYSYGNVAPL	300
QY	301	NNGKTLFLNNVASPVYIAAKOPTSGOASNTNNYDGGAIKCKGAQ-AGSNNSSVSFPD	359
Db	301	NNGKTLFLNNVASPVYIAAKOPTSGOASNTNNYDGGAIKCKGAQ-AGSNNSSVSFPD	360
QY	360	GEVVVFFSSNVAAGKGGAIYAKLSVANCGPVQFLRNIANDGGAIYLGESGELSADYG	419
Db	361	GEVVVFFSSNVAAGKGGAIYAKLSVANCGPVQFLRNIANDGGAIYLGESGELSADYG	420
QY	420	DIIFDGNLKRTAKENAADVNGVTSSQAI SMGSGKITTLRAKAGHQLFNDPIEWANG	479
Db	421	DIIFDGNLKRTAKENAADVNGVTSSQAI SMGSGKITTLRAKAGHQLFNDPIEWANG	480
QY	480	NOPAQSSKLLKINDGEGYTDIVFANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSQTCG	539
Db	481	NOPAQSSKLLKINDGEGYTDIVFANGSSTLYQNTVIEQGRIVLREKAKLSVNSLSQTCG	540
QY	540	SLYMEAGSTWDFVTPQQPPQPPAANQLITLSNLHLSSLANNATVTPPTPPAQDSHP	599
Db	541	SLYMEAGSTWDFVTPQQPPQPPAANQLITLSNLHLSSLANNATVTPPTPPAQDSHP	600
QY	600	AVIGSTTAGSVTISGPIPEDLDDTAYDRYDWLGSNOKINVLKLGKYPANASDLTL	659
Db	601	AVIGSTTAGSVTISGPIPEDLDDTAYDRYDWLGSNOKINVLKLGKYPANASDLTL	660
QY	660	GNEMPKYGYQGSWKLAWDPNTANNPYTLKATWTKTYNPGPVERVASLVPNSLWSILDI	719
Db	661	GNEMPKYGYQGSWKLAWDPNTANNPYTLKATWTKTYNPGPVERVASLVPNSLWSILDI	720
QY	720	RSASAIQASVDGRSYCRGLWVSGVSNPFYHDDALGQGYRISGYSIGANSYFGSSMF	779
Db	721	RSASAIQASVDGRSYCRGLWVSGVSNPFYHDDALGQGYRISGYSIGANSYFGSSMF	780
QY	780	GLAPTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYGFGNQHMT	839
Db	781	GLAPTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYGFGNQHMT	840
QY	840	SYTFAESDVRDNNCLAGEIGAGLPVITPSKLYLNEIRPFVQAEFSYADHESFTEGD	899
Db	841	SYTFAESDVRDNNCLAGEIGAGLPVITPSKLYLNEIRPFVQAEFSYADHESFTEGD	900
QY	900	QARAFKSGHLLNLSPVGVKDFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	959
Db	901	QARAFKSGHLLNLSPVGVKDFDRCSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW	960
QY	960	TTDAFHARHGIVVRGSMYASLTSNIEVYGHGRYEDRSRGYGLSAGSRVRF	1012
Db	961	TTDAFHARHGIVVRGSMYASLTSNIEVYGHGRYEDRSRGYGLSAGSKVRF	1013
RESULT 22			
ID	AAV16737		
XX	AAV16737	standard; protein; 1013 AA.	
XX	AAV16737;		
XX	21-JUL-1999	(first entry)	
XX	C. trachomatis B serovar HMW protein.		

PT New Chlamydia protein useful for treating conjunctivitis, urethritis and
PT cervical cancer.
XX Claim 4; Page 119-123; 141pp; English.
XX The invention relates to an isolated Chlamydia species high molecular
CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
CC be used for preventing, treating or ameliorating a disorder related to
CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
CC cancer, infertility, arteriosclerosis and atherosclerosis. The products
CC can also be used for detection and diagnosis. The present sequence
CC represents a C. trachomatis HMW protein
XX
SQ Sequence 1013 AA;

Query Match 97.0%; Score 5106.5; DB 2; Length 1013;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 980; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

QY 1 MQTSFKFPLSLMILAYSCSLGGGYAAEIMVPOGIYDGETLTVSPVTVIGDPSGTTVF 60
DB 1 MQTSFKFPLSLMILAYSCSLGGGYAAEIMVPOGIYDGETLTVSPVTVIGDPSGTTVF 60

QY 61 SAGELTAKNLSIAALPLSCFNLGSPVTLGRGHSITFENIRTSNGAALSNAADGL 120
DB 61 SAGELTAKNLSIAALPLSCFNLGSPVTLGRGHSITFENIRTSNGAALSNAADGL 120

QY 121 FTIEGFKELFNCNLSIAALPLSCFNLGSPVTLGRGHSITFENIRTSNGAALSNAADGL 180
DB 121 FTIEGFKELFNCNLSIAALPLSCFNLGSPVTLGRGHSITFENIRTSNGAALSNAADGL 180

QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGACQVTVTSFSAANEAPAFVAN 240
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGACQVTVTSFSAANEAPAFVAN 240

QY 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVFDFGNVAVRGVGIYSYGNVAPL 300
DB 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVSFRNTAVFDFGNVAVRGVGIYSYGNVAPL 300

QY 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDDGAI FCKNGAQ-AGSNNSGSVDF 359
DB 301 NNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDDGAI FCKNGAQ-AGSNNSGSVDF 359

QY 360 GEGWPFSSNVAAGKGAIYAKLSVANGCPVQFLNANDGGAIVLGSSELISADYG 419
DB 360 GEGWPFSSNVAAGKGAIYAKLSVANGCPVQFLNANDGGAIVLGSSELISADYG 419

QY 420 DIIFDGNLAKRTAKENADVNGVTSSQAISMGSGGKIITLRAKAGHOILFNDDPIEMANGN 479
DB 420 DIIFDGNLAKRTAKENADVNGVTSSQAISMGSGGKIITLRAKAGHOILFNDDPIEMANGN 479

QY 480 NQPAQSKLLKINDGGYTGDIYFANGSSTLYQNVTEQGRIVLRKAKLSVNSLSQTCG 539
DB 480 NQPAQSKLLKINDGGYTGDIYFANGSSTLYQNVTEQGRIVLRKAKLSVNSLSQTCG 539

QY 540 SLYMEAGSTWDFVTPPOPPQPPAANOLIITLSNLHLSLSLLANNVTPNPPADQSH 599
DB 540 SLYMEAGSTWDFVTPPOPPQPPAANOLIITLSNLHLSLSLLANNVTPNPPADQSH 599

QY 541 SLYMEAGSTWDFVTPPOPPQPPAANOLIITLSNLHLSLSLLANNVTPNPPADQSH 600
DB 541 SLYMEAGSTWDFVTPPOPPQPPAANOLIITLSNLHLSLSLLANNVTPNPPADQSH 600

QY 600 AVIGSTTAGSVTISGRTFFEDLDDTAYDRYDMLGSKNQKINVLKQLGTPPPANAPSDTL 659
DB 600 AVIGSTTAGSVTISGRTFFEDLDDTAYDRYDMLGSKNQKINVLKQLGTPPPANAPSDTL 659

QY 660 GNEMPKYGYQGSWKLAWDPNTANNPPTLTKATWTKGYNPGERVASLVNSLWGLSILDI 719
DB 660 GNEMPKYGYQGSWKLAWDPNTANNPPTLTKATWTKGYNPGERVASLVNSLWGLSILDI 719

QY 720 RSAHSAIQASVGRSVCRLGWGVSNSFFVHDDALGQGYRIYSGGYSLCANSYFGSSMF 779
DB 720 RSAHSAIQASVGRSVCRLGWGVSNSFFVHDDALGQGYRIYSGGYSLCANSYFGSSMF 779

QY 721 RSAHSAIQASVGRSVCRLGWGVSNSFFVHDDALGQGYRIYSGGYSLCANSYFGSSMF 780
DB 721 RSAHSAIQASVGRSVCRLGWGVSNSFFVHDDALGQGYRIYSGGYSLCANSYFGSSMF 780

QY 780 GLAFTEVGRSKDYVWCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASYGFGNQHMKT 839
DB 780 GLAFTEVGRSKDYVWCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASYGFGNQHMKT 840

QY 840 SYTPAESDVRWNNCLAGIEAGLPIVITPSKUYLNLRLPFPVQAEFSYADHESFTEGD 899
DB 840 SYTPAESDVRWNNCLAGIEAGLPIVITPSKUYLNLRLPFPVQAEFSYADHESFTEGD 900

QY 900 QARAFKSHLLNLSPVGVKFDRCSSHPNPKYSFMAAYICDAYRTISGTETLLSHQETW 959
DB 900 QARAFKSHLLNLSPVGVKFDRCSSHPNPKYSFMAAYICDAYRTISGTETLLSHQETW 960

QY 960 TTDAFLHARHGVVGRGSMYASLTNSIEVYHGRYEDRSDASRGYGLSAGSRVRF 1012
DB 960 TTDAFLHARHGVVGRGSMYASLTNSIEVYHGRYEDRSDASRGYGLSAGSRVRF 1013

RESULT 24
AAG83207
ID AAG83207 standard; protein; 1006 AA.
XX
AC AAG83207;
DT 05-SEP-2001 (first entry)
XX
DE Protein encoded by Chlamydia trachomatis pmpG gene.
XX
KW Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Cap1; CT529; OMCB;
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
OS Chlamydia trachomatis.
XX
PN WO200140474-A2.
XX
PD 07-JUN-2001.
XX
PF 04-DEC-2000; 2000WO-US032919.
XX
PR 03-DEC-1999; 99US-00454684.
PR 19-APR-2000; 2000US-00556877.
PR 20-JUN-2000; 2000US-00598419.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAM, Fling SP, Scholler J;
XX
DR WPI; 2001-374831/39.
XX
PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease.
XX
PS Claim 2; Page 214-216; 295pp; English.
XX
PS The present sequence is provided in a specification relating to compounds
CC and methods for the treatment and diagnosis of chlamydial infection. The
CC compounds provided include polypeptides and fusion proteins comprising
CC immunogenic portions of Chlamydia antigens and DNA sequences encoding
CC such polypeptides. They are useful for vaccinating against chlamydial
CC infection, which causes pelvic inflammatory disease, trachoma, acute
CC respiratory tract infections, atherosclerosis and heart disease
XX
SQ Sequence 1006 AA;

Query Match 96.6%; Score 5090; DB 4; Length 1006;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 MVFPGIYDGETLTVSPVTVIGDPSGTTVFSAGELTKNLDNSIAALPLSCFNLGSGFT 90
DB 31 MVFPGIYDGETLTVSPVTVIGDPSGTTVFSAGELTKNLDNSIAALPLSCFNLGSGFT 90

Db 25 MIPQGIYDGETLTVSPPYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFGLNLSFT 84
Qy 91 VLGRGHSLTPTENIRTSNGAALSAAADGLFTIEGFKELSFNSCNSLLAVLPAATTNKS 150
Db 85 VLGRGHSLTPTENIRTSNGAALSAAADGLFTIEGFKELSFNSCNSLLAVLPAATTNKS 144
Qy 151 QTPTTSTSPNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOE 210
Db 145 QTPTTSTSPNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOE 204
Qy 211 NTAQADGGACQVTSFSSAMANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 270
Db 205 NTAQADGGACQVTSFSSAMANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 264
Qy 271 SPSRNTAVFEDGNVARVGGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 330
Db 265 SPSRNTAVFEDGNVARVGGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 324
Qy 331 SNNTYDGGAIFFCKNGAQAGSNNSGSVSPDGRGVVPFSSNVAAGKGAIYAKKLSVANC 390
Db 325 SNNTYDGGAIFFCKNGAQAGSNNSGSVSPDGRGVVPFSSNVAAGKGAIYAKKLSVANC 384
Qy 391 VOFLRNANDGGAIYLGESBELSLADYGDIIFDGNLAKRTAKENAADVNGVTVSSQALSM 450
Db 385 VOFLRNANDGGAIYLGESBELSLADYGDIIFDGNLAKRTAKENAADVNGVTVSSQALSM 444
Qy 451 GSGGKIITLRAKAGHQLFNDPIEMANGNPPAQSSKLLKINDGEYTGDIIVFANGSSTL 510
Db 445 GSGGKIITLRAKAGHQLFNDPIEMANGNPPAQSSKLLKINDGEYTGDIIVFANGSSTL 504
Qy 511 YQNTVIEGRIVLREKAKLSVNSISOTGGSILYMEAGSTMDPVTPOPPQAPAAQLITLS 570
Db 505 YQNTVIEGRIVLREKAKLSVNSISOTGGSILYMEAGSTMDPVTPOPPQAPAAQLITLS 564
Qy 571 NLHLSLSLLANNAVTPPTPPAQDSHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDRYD 630
Db 565 NLHLSLSLLANNAVTPPTPPAQDSHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDRYD 624
Qy 631 WLGNQKINVLKQLGTPKPNASBDLTLGNEMPKYGYQGSWKLAWDPTANNPYTLKA 690
Db 625 WLGNQKINVLKQLGTPKPNASBDLTLGNEMPKYGYQGSWKLAWDPTANNPYTLKA 684
Qy 691 TWTKTGNRGPERSVASLVPNSLMSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNPFYH 750
Db 685 TWTKTGNRGPERSVASLVPNSLMSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNPFYH 744
Qy 751 DRDALGQGYRIYSGGYSILGANSYFGSSMFLAFTFVGRSKDYVVCRSNHACIGSVYLS 810
Db 745 DRDALGQGYRIYSGGYSILGANSYFGSSMFLAFTFVGRSKDYVVCRSNHACIGSVYLS 804
Qy 811 TQOALCGSYLPDGAFTIRASYFGNQHMKTSTYTPAEESDVRDNNCLAGEIGAGLPVITP 870
Db 805 TQOALCGSYLPDGAFTIRASYFGNQHMKTSTYTPAEESDVRDNNCLAGEIGAGLPVITP 864
Qy 871 SKLYNELRPVQAEPSYADHESFTEBDQARAFKSGHLLNLSVPVGVKPRDCSSTHPNK 930
Db 865 SKLYNELRPVQAEPSYADHESFTEBDQARAFKSGHLLNLSVPVGVKPRDCSSTHPNK 924
Qy 931 YSFMAAYICDAYRTISGTETTLTLLSHQETWTWTDAPHARHGTVVVGSMYASLTSNIEVGH 990
Db 925 YSFMAAYICDAYRTISGTETTLTLLSHQETWTWTDAPHARHGTVVVGSMYASLTSNIEVGH 984
Qy 991 GRYEYRDASRGYGLSAGSRVRF 1012
Db 985 GRYEYRDASRGYGLSAGSRVRF 1006

RESULT 25

ABB94178

ID ABB94178 standard; protein; 1006 AA.

XX

AC ABB94178;

XX

DT 05-JUN-2002 (first entry)
XX Chlamydia protein sequence SEQ ID NO:190.
DE Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;
XX antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response.
KW Chlamydia sp.
OS
XX
FN WO200208267-A2.
XX
PD 31-JAN-2002.
XX
PP 20-JUL-2001; 2001WO-US023121.
XX
PP 20-JUL-2001; 2000US-00620412.
XX
PR 23-APR-2001; 2001US-00841132.
XX
XX (CORI-) CORIXA CORP.
XX
XX Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX MPI; 2002-179901/23.
XX
XX Novel compositions comprising Chlamydia Cap1 protein and its use in the
XX treatment of Chlamydia infection.
XX
XX Example 1; Page 245-247; 537pp; English.
XX
XX The present invention describes compositions comprising a Chlamydia Cap1
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX vaccines. Compounds from the present invention can be used for eliciting
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment of a Chlamydia
XX infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
XX sequences used in the exemplification of the present invention
XX
SQ Sequence 1006 AA;
Query Match 96.6%; Score 5090; DB 5; Length 1006;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 979; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 31 MIPQGIYDGETLTVSPPYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFGLNLSFT 90
Db 25 MIPQGIYDGETLTVSPPYTVIGDPSTTVFSAGELTLKNLDSIAALPLSCFGLNLSFT 84
Qy 91 VLGRGHSLTPTENIRTSNGAALSAAADGLFTIEGFKELSFNSCNSLLAVLPAATTNKS 150
Db 85 VLGRGHSLTPTENIRTSNGAALSAAADGLFTIEGFKELSFNSCNSLLAVLPAATTNKS 144
Qy 151 QTPTTSTSPNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOE 210
Db 145 QTPTTSTSPNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOE 204
Qy 211 NTAQADGGACQVTSFSSAMANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 270
Db 205 NTAQADGGACQVTSFSSAMANEAPAFVANVAVRGGGIAAVQDQGGVSSSTSTEDPVV 264
Qy 271 SPSRNTAVFEDGNVARVGGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 330
Db 265 SPSRNTAVFEDGNVARVGGGIYSYGNVAFLLNGKTLFLNNVASPVYIAAKQPTSGQASNT 324
Qy 331 SNNTYDGGAIFFCKNGAQAGSNNSGSVSPDGRGVVPFSSNVAAGKGAIYAKKLSVANC 390
Db 325 SNNTYDGGAIFFCKNGAQAGSNNSGSVSPDGRGVVPFSSNVAAGKGAIYAKKLSVANC 384

QY 391 VQFLRNANDGGAIYLGSGELSLADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450
DB 385 VQFLRNANDGGAIYLGSGELSLADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 444
QY 451 GSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
DB 445 GSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 504
QY 511 YQNVITIEQRIIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570
DB 505 YQNVITIEQRIIVLREKAKLSVNSLSQSGSLYMEAGSTLDFVTPPOPPQPPAANQLITLS 564
QY 571 NLHLSLSLLANNVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYDRYD 630
DB 565 NLHLSLSLLANNVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYDRYD 624
QY 631 WLGSNQKINVLKQLGTPKANAPSDLTILGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 690
DB 625 WLGSNQKINVLKQLGTPKANAPSDLTILGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 684
QY 691 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFVH 750
DB 685 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFVH 744
QY 751 DRDALGQGYRYISGGYSLGANSYFGSSMFLGAPTEVFGRSKDYVWCRSNHHACIGSVYLS 810
DB 745 DRDALGQGYRYISGGYSLGANSYFGSSMFLGAPTEVFGRSKDYVWCRSNHHACIGSVYLS 804
QY 811 TQOALCGSYLFGDAFIRASVFGNQHMKTSYTPAEBSDVRWNNCLAGHIGAGLPVITPP 870
DB 805 TQOALCGSYLFGDAFIRASVFGNQHMKTSYTPAEBSDVRWNNCLAGHIGAGLPVITPP 864
QY 871 SKLYLNELPFPVQAEFSYADHESFTEGDOARAFKSGHLLNLSVPYGVKFCDCSSSTHPNK 930
DB 865 SKLYLNELPFPVQAEFSYADHESFTEGDOARAFKSGHLLNLSVPYGVKFCDCSSSTHPNK 924
QY 931 YSPMAAYICDVAETISGTETLLSHOETWTTDAPHLARHGWWVRGSMYASLTSNIEVYGH 990
DB 925 YSPMAAYICDVAETISGTETLLSHOETWTTDAPHLARHGWWVRGSMYASLTSNIEVYGH 984
QY 991 GRYEYRDASRGYLSAGSRVRF 1012
DB 985 GRYEYRDASRGYLSAGSRVRF 1006
RESULT 26
AAB13633
ID AAB13633 standard; protein; 982 AA.
XX AAB13633;
AC AAB13633;
DT 02-FEB-2001 (first entry)
XX C. trachomatis pmpG gene protein.
DE XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease; antibacterial.
XX
OS Chlamydia trachomatis.
XX
FH Key Location/Qualifiers
FT Misc-difference 981
FT /note= "Unspecified amino acid"
FX
FN WO200034483-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-US029012.
XX
PR 08-DEC-1998; 98US-00208277.

PR 08-APR-1999; 99US-00288594.
PR 01-OCT-1999; 99US-00410568.
PR 22-OCT-1999; 99US-00426571.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skelky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
PT acid sequence encoded by polynucleotide sequence.
XX
PS Claim 2; Page 181-184; 256pp; English.
XX
CC The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamydiae are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention
XX
SQ Sequence 982 AA;
Query Match 96.5%; Score 5084; DB 3; Length 982;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 31 MYPQGIYDGETLTVSPFTYVIGDPSTTVFSAGELTLKNDNSIAALPLSCFNLGSPFT 90
DB 1 MYPQGIYDGETLTVSPFTYVIGDPSTTVFSAGELTLKNDNSIAALPLSCFNLGSPFT 60
QY 91 VLGRGHSILTFENIRTSNGAALSNAAADGLFTIEGFKELSFNCNSLLAVLPAATNKG 150
DB 61 VLGRGHSILTFENIRTSNGAALSNAAADGLFTIEGFKELSFNCNSLLAVLPAATNKG 120
QY 151 QTPPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 210
DB 121 QTPPTTSTPSNGTIYSKTDLLLNNEKFSFYNLVSGDGAIDAKSLTVQGISKLCVFOE 180
QY 211 NTAQADGGACQVVTSPSAMANEAPAFVANVAGVRGGIAAVQDGGQGVSSSTSTEDPVV 270
DB 181 NTAQADGGACQVVTSPSAMANEAPAFVANVAGVRGGIAAVQDGGQGVSSSTSTEDPVV 240
QY 271 SFSRNTAVFDPGNVAVRGVGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASNT 330
DB 241 SFSRNTAVFDPGNVAVRGVGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASNT 300
QY 331 SNNYGCGAIFCKNGAQAQSNNSGVSFDGEGVVFSSNVVAGKGAIVAKKLSVANC 390
DB 301 SNNYGCGAIFCKNGAQAQSNNSGVSFDGEGVVFSSNVVAGKGAIVAKKLSVANC 360
QY 391 VQFLRNANDGGAIYLGSGELSLADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450
DB 361 VQFLRNANDGGAIYLGSGELSLADYGDIIIFDGNLKRKTAKENAADVNGVTVSSQAISM 420
QY 451 GSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
DB 421 GSGGKITTLRAKAGHQILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIIVFANGSSTL 480
QY 511 YQNVITIEQRIIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570
DB 481 YQNVITIEQRIIVLREKAKLSVNSLSQSGSLYMEAGSTLDFVTPPOPPQPPAANQLITLS 540

Qy	571	NLHLSJSSLIANNVNTPTNPQAQDHPAVIGSTTAGSVTISGPIFFEDLDTDAYDRYD	630
Db	541	NLHLSJSSLIANNVNTPTNPQAQDHPAVIGSTTAGSVTISGPIFFEDLDTDAYDRYD	600
Qy	631	WLGSNOKINVLKQLGCTKPANAPASDLTLGNEMPKYGQGSWKLANDPNTANNPPTLKA	690
Db	601	WLGSNOKINVLKQLGCTKPANAPASDLTLGNEMPKYGQGSWKLANDPNTANNPPTLKA	660
Qy	691	TWTKTGYNPGPERVASLVPSNLWSGISILDIRSAHSALQASVDGRSYCKRGLMVGVSNNFFYH	750
Db	661	TWTKTGYNPGPERVASLVPSNLWSGISILDIRSAHSALQASVDGRSYCKRGLMVGVSNNFFYH	720
Qy	751	DRDALCOGGRYIISGGYSILGANYSFOSSNFGFLAFTVEVGRSKDYVCBSNHACIGSVYLS	810
Db	721	DRDALCOGGRYIISGGYSILGANYSFOSSNFGFLAFTVEVGRSKDYVCBSNHACIGSVYLS	780
Qy	811	TQOALCGSLFQDAPIRASYGFGNOHKMTSYTFABESDVWRDNCLAGELGAGLPVTVP	870
Db	781	TQOALCGSLFQDAPIRASYGFGNOHKMTSYTFABESDVWRDNCLAGELGAGLPVTVP	840
Qy	871	SKLYLNELAPPVOAFSPSADHESPTEEGDOARAFKSGHLNLSVPVGVKFPDRCSSTHPNK	930
Db	841	SKLYLNELAPPVOAFSPSADHESPTEEGDOARAFKSGHLNLSVPVGVKFPDRCSSTHPNK	900
Qy	931	YSPMAAYICDAYRTISGTETILLSHQETWTTDDAFILAHGVVVRGSMYASILTNIENVYGH	990
Db	901	YSPMAAYICDAYRTISGTETILLSHQETWTTDDAFILAHGVVVRGSMYASILTNIENVYGH	960
Qy	991	GRYEYDASRGYGLSAGSRVRF	1012
Db	961	GRYEYDASRGYGLSAGSKVXF	982

RESULT 27

RESULTS 2/
AAG83201

AA83201
ID AA83201 standard; protein; 982 AA.

XX
TACCOMI

AC AAG83201;

XX

DT 05-SEP-2001 (first entry)

XX

DE Protein encoded by *Chlamydia trachomatis* pmpG gene.
yy

XX
MA
XX

Chumbe vaccine: infection: function protection: anti-gen.

KW Chlamydia; vaccine; inflammatory

acute respiratory distress syndrome

[illegible]

XX
FOLIO 100
22

OS Chlamydia trachomatis.

XX

PN WO200140474-A2:

XX PD 07-TIN-2001

FD 07-JUN-2001.
XX

04-DRC-2000: 2000WO-

FF XX
04-DEC-2000; 2000MO-05052315.

PR 03-DEC-1999; 99US-00454684.

PR 19-APR-2000; 2000UG-00556877.

PR 20-JUN-2000; 2000US-00598419.

XX
 23
 /CONT / CONTY3 CONT3

and methods for the treatment and diagnosis of chlamydial infection. The compounds provided include polypeptides and fusion proteins comprising immunogenic portions of Chlamydia antigens and DNA sequences encoding such polypeptides. They are useful for vaccinating against chlamydial infection, which causes pelvic inflammatory disease, trachoma, acute respiratory tract infections, atherosclerosis and heart disease

Query Match	96.5%	Score 5084	DB 4	Length 982
Best Local Similarity	99.6%	Pred. No. 0		
Matches 978; Conservative	2	Mismatches	2	Indels 0; Gaps 0;

Qy	31	MVPGIYDGETLTVSPYVTVYGDPSGTTVFSAGELTLKNLNSIAALPIASCFONLGSGFT	90
Db	1	MI PQGIYDGETLTVSPYVTVYGDPSGTTVFSAGELTLKNLNSIAALPIASCFONLGSGFT	60
Qy	91	VLGRGHSLTPEINTSTNGAALSNSAADGLFTIEGFKELSFNSNCNSLLAVLPAATTNKG	150
Db	61	VLGRGHSLTPEINTSTNGAALSNSAADGLFTIEGFKELSFNSNCNSLLAVLPAATTNKG	120
Qy	151	QTPPTTSTPSNGTYSKNTDLLLNNEKFPFYSNLVSGDGAIDAKSLTVOGISKL CVFOE	210
Db	121	QTPPTTSTPSNGTYSKNTDLLLNNEKFPFYSNLVSGDGAIDAKSLTVOGISKL CVFOE	180
Qy	211	NTAQADGACQVVTSPSAMNARAPIAVANVAGVGGITAAVODGOGVSSBSTSTEDPVV	270
Db	181	NTAQADGACQVVTSPSAMNARAPIAVANVAGVGGITAAVODGOGVSSBSTSTEDPVV	240
Qy	271	SFSRNTAVFEDGNVAVGGIYISYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGQASNT	330
Db	241	SFSRNTAVFEDGNVAVGGIYISYGNVAFPLNNGKTLFLNNVASPVYIAAKQPTSGQASNT	300
Qy	331	SNNYGDGGAIFCKNGAQAGSNNSGSVSPQEGVVFPSSNVVAGKGGAIYAKKLSVANCGP	390
Db	301	SNNYGDGGAIFCKNGAQAGSNNSGSVSPQEGVVFPSSNVVAGKGGAIYAKKLSVANCGP	360
Qy	391	VQFLRNIANDGGAIYLGESEGLSLSADYDGIIPDGNLKTAKENAADVNGVTVSSQAIISM	450
Db	361	VQFLRNIANDGGAIYLGESEGLSLSADYDGIIPDGNLKTAKENAADVNGVTVSSQAIISM	420
Qy	451	GSQGIKITTILAKAGHOILFNDPIEMANGNQPQAQSSKLLKINDGEGYTGDIIVFANGSSSTL	510
Db	421	GSQGIKITTILAKAGHOILFNDPIEMANGNQPQAQSSKLLKINDGEGYTGDIIVFANGSSSTL	480
Qy	511	YQNVTTIEQGRILVREKAKLSVNSLSTQGSGLYMEAGSTWDFVTPQPPQPPAANQILITUS	570
Db	481	YQNVTTIEQGRILVREKAKLSVNSLSTQGSGLYMEAGSTWDFVTPQPPQPPAANQILITUS	540
Qy	571	NLHLSLSLLANNAVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYRD	630
Db	541	NLHLSLSLLANNAVTPNTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDLDTAYRD	600
Qy	631	WLGSNQKINVLKLOLGTKPPANAPDLTIGNEMPKYCYGGSWKLAWDPTANNPPTLKA	690
Db	601	WLGSNQKINVLKLOLGTKPPANAPDLTIGNEMPKYCYGGSWKLAWDPTANNPPTLKA	660
Qy	691	TWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	750
Db	661	TWTKTGNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYH	720
Qy	751	DRDALGGYRYISGGYVSLGANSYFGSSMPGLAFTVEVFGRSKDYVVCRSNHHACIGSVYLS	810
Db	721	DRDALGGYRYISGGYVSLGANSYFGSSMPGLAFTVEVFGRSKDYVVCRSNHHACIGSVYLS	780
Qy	811	TQOALCGSYLFGDAFIRASVFGNQHMKTSYTPAESDVRWNNCLAGIBGLP LVTI TP	870
Db	781	TQOALCGSYLFGDAFIRASVFGNQHMKTSYTPAESDVRWNNCLAGIBGLP LVTI TP	840
Qy	871	SKLYTANELRPFVQAFYSADHESFTEEGDQARFAPSGLHLLNTSVPVGVKFDRCSSHTPNK	930
Db	841	SKLYTANELRPFVQAFYSADHESFTEEGDQARFAPSGLHLLNTSVPVGVKFDRCSSHTPNK	900

QY 931 YSFMAAICDAVRTISGTTTLLSHQETWTTDAFHLARHGVVVRGSMVASLTSNIEVYGH 990
 DB 901 YSFMAAICDAVRTISGTTTLLSHQETWTTDAFHLARHGVVVRGSMVASLTSNIEVYGH 960
 QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
 DB 961 GRYEYRDASRGYGLSAGSKVXF 982

RESULT 28
 ABB94172
 ID ABB94172 standard; protein; 982 AA.
 XX
 AC ABB94172;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Chlamydia protein sequence SEQ ID NO:176.
 XX
 KW Chlamydia infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.
 XX
 OS Chlamydia sp.
 XX
 PN WO200208267-A2.
 XX
 PD 31-JAN-2002.
 XX
 PP 20-JUL-2001; 2001WO-US023121.
 XX
 PR 20-JUL-2000; 2000US-00620412.
 PR 23-APR-2001; 2001US-00841132.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Fling SP, Skeiky YAW, Probst P, Bhatia A;
 XX
 DR WPI; 2002-179901/23.
 XX
 PT Novel compositions comprising Chlamydia Cap1 protein and its use in the
 PT treatment of Chlamydia infection.
 XX
 PS Disclosure; Page 221-223; 537pp; English.
 XX
 CC The present invention describes compositions comprising a Chlamydia Cap1
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.
 CC Chlamydia DNA and protein sequences from the present invention can have
 CC antibacterial and immunostimulant activities, and can be used in
 CC vaccines. Compounds from the present invention can be used for eliciting
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell
 CC response or inhibiting the development of a Chlamydia infection in an
 CC animal. Methods from the present invention can be used for detecting the
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia
 CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 CC sequences used in the exemplification of the present invention
 XX
 SQ Sequence 982 AA;

Query Match 96.5%; Score 5084; DB 5; Length 982;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 31 MIPQGIYDGETLVSFPYTVIGDPSTVFSAGBELTLKNDNSIAALPLSCFNGLLGSFT 90
 DB 1 MIPQGIYDGETLVSFPYTVIGDPSTVFSAGBELTLKNDNSIAALPLSCFNGLLGSFT 60
 QY 91 VLGRGHSLLTENIRTSNGAALSNSAADGLFTIEGPKELSFNSCNSLLAVLPAATTNKG 150
 DB 61 VLGRGHSLLTENIRTSNGAALSNSAADGLFTIEGPKELSFNSCNSLLAVLPAATTNKG 120
 QY 151 QTPPTTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVQFE 210

DB 121 QTPPTTSPNGTIYSKTDLLLNNEKFSFYSNLVSGDGAIDAKSLTVQGISKLCVQFE 180
 QY 211 NTAQADGGACQCVVTSFSAANEAPAFVANVAGVRGGIAAIVQDQCGQCVSSSTSTEDPV 270
 DB 181 NTAQADGGACQCVVTSFSAANEAPAFVANVAGVRGGIAAIVQDQCGQCVSSSTSTEDPV 240
 QY 271 SFSRNTAVEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSOASNT 330
 DB 241 SFSRNTAVEFDGNVARVGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSOASNT 300
 QY 331 SNNYGDGGAIFCKGGAQAGSNNSGVSFDPGCVVPFSSNVAAGKGAIAKLSVANGCP 390
 DB 301 SNNYGDGGAIFCKGGAQAGSNNSGVSFDPGCVVPFSSNVAAGKGAIAKLSVANGCP 360
 QY 391 VOFLRNANDGGAIVLGSSELSSADYGDIIFDGNLKRKTAKENAADVNGVTVSSQAISM 450
 DB 361 VOFLRNANDGGAIVLGSSELSSADYGDIIFDGNLKRKTAKENAADVNGVTVSSQAISM 420
 QY 451 GSGGKITTLRAKAGHQIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIYFANGSSTL 510
 DB 421 GSGGKITTLRAKAGHQIILFNDPIEMANGNNOQAQSSKLLKINDGEGYTGDIYFANGSSTL 480
 QY 511 YONVTIEQGRIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPPOPPQPPAANQLITLS 570
 DB 481 YONVTIEQGRIVLREKAKLSVNSLSQTGSLYMEAGSTLDFVTPPOPPQPPAANQLITLS 540
 QY 571 NLHLSLSLLANNVTPNPPADQSHPAVIGSTTAGSVTISGPIFPEDLDODTAYDRYD 630
 DB 541 NLHLSLSLLANNVTPNPPADQSHPAVIGSTTAGSVTISGPIFPEDLDODTAYDRYD 600
 QY 631 WLGSNOKINVLKQLGTKPANAPSDLTIGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 690
 DB 601 WLGSNOKINVLKQLGTKPANAPSDLTIGNEMPKYGYQGSWKLAWDPNTANNPPTLKA 660
 QY 691 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSNFFYH 750
 DB 661 TWTKTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSNFFYH 720
 QY 751 DRDALCGYRYISGGYSLGANSYRGSSMFGLAFTVEVGRSKDYVVCRSNHHACIGSVYLS 810
 DB 721 DRDALCGYRYISGGYSLGANSYRGSSMFGLAFTVEVGRSKDYVVCRSNHHACIGSVYLS 780
 QY 811 TQOALCGSYLFGDAFIRASYGFGNQHMKTSYTFABESDVRDNNCLAGEIGAGLPVITP 870
 DB 781 TQOALCGSYLFGDAFIRASYGFGNQHMKTSYTFABESDVRDNNCLAGEIGAGLPVITP 840
 QY 871 SKLYLNLAPFFVQAFSADHESFTEEGDQAFKSGHLLNLSPVGVKVFDRCSSTHPNK 930
 DB 841 SKLYLNLAPFFVQAFSADHESFTEEGDQAFKSGHLLNLSPVGVKVFDRCSSTHPNK 900
 QY 931 YSFMAAICDAVRTISGTTTLLSHQETWTTDAFHLARHGVVVRGSMVASLTSNIEVYGH 990
 DB 901 YSFMAAICDAVRTISGTTTLLSHQETWTTDAFHLARHGVVVRGSMVASLTSNIEVYGH 960
 QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
 DB 961 GRYEYRDASRGYGLSAGSKVXF 982

RESULT 29

AAB13639
 ID AAB13639 standard; protein; 1006 AA.

XX
 AC AAB13639;

DT 02-FEB-2001 (first entry)

XX C. trachomatis pmpG gene protein.

XX Chlamydial infection; sexually transmitted disease;

KW pelvic inflammatory disease; PID; tubal obstruction; infertility;

KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;

KW coronary heart disease; antibacterial.
 XX Chlamydia trachomatis.
 OS WO200034483-A2.
 PN 15-JUN-2000.
 PD 08-DEC-1999; 99WO-US029012.
 PF 08-DEC-1998; 98US-00208277.
 XX 08-APR-1999; 99US-00288594.
 PR 01-OCT-1999; 99US-00410568.
 PR 22-OCT-1999; 99US-00426571.
 XX (CORI-) CORIXA CORP.
 PA Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
 PI WPI, 2000-431303/37.
 DR Isolated polypeptide for diagnosis and treatment of Chlamydia infection
 XX comprises immunogenic portion of Chlamydia antigen, which comprises amino
 PT acid sequence encoded by polynucleotide sequence.
 PT Claim 2; Page 208-210; 256pp; English.
 XX The present invention relates to new nucleic acid sequences and the
 CC proteins encoded by the nucleic acid sequences. The encoded proteins
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia
 CC infection. Chlamydiae are intracellular bacterial pathogens that are
 CC responsible for a wide variety of human infections. C. trachomatis
 CC infection is one of the most common sexually transmitted diseases and
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is
 CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
 CC major cause of acute respiratory tract infections in humans and is also
 CC thought to play a role in the pathogenesis of atherosclerosis and
 CC coronary heart disease. The present sequence is a protein isolated in the
 CC present invention
 XX Sequence 1006 AA;
 BQ

Query Match 96.5%; Score 5083; DB 3; Length 1006;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 978; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 31 MVRQGIYDGETLTWSPPYTVIGDPSGTTVPBAGELTLKNLDSNLAALPLSCFNLGSPT 90
 DB 25 MIPQGIYDGETLTWSPPYTVIGDPSGTTVPBAGELTLKNLDSNLAALPLSCFNLGSPT 84
 QY 91 VLGRGHSITFNIRITSTNGAALSNSAADLFTIEGPKELSPNCNSLLAVLPAATTNKG 150
 DB 85 VLGRGHSITFNIRITSTNGAALSNSAADLFTIEGPKELSPNCNSLLAVLPAATTNKG 144
 QY 151 QTPTTTSPSNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTVQGISKLCVFOE 210
 DB 145 QTPTTTSPSNGTIYSKTDLLLNNEKSPFYSNLVSGDGAIDAKSLTVQGISKLCVFOE 204
 QY 211 NTAQADGACQVWTSFSAANEAFIAFVANVAGVGGIAAVDQGGVSSSTSTEDPVV 270
 DB 205 NTAQADGACQVWTSFSAANEAFIAFVANVAGVGGIAAVDQGGVSSSTSTEDPVV 264
 QY 271 SFSRNTAVEPDGNVARGGGIYSYGNVAFNNGKTLFLNNVASPYTAAKOPTSGQASNT 330
 DB 265 SFSRNTAVEPDGNVARGGGIYSYGNVAFNNGKTLFLNNVASPYTAAKOPTSGQASNT 324
 QY 331 SNNTYGDGGAIFCKXGAQAGSNNSGVSFDPGEGVVPFSSNVAAGKGAIAKLSVANCGP 390
 DB 325 SNNTYGDGGAIFCKXGAQAGSNNSGVSFDPGEGVVPFSSNVAAGKGAIAKLSVANCGP 384
 QY 391 VOFLRNITANDGGAIIYLGESGELSADYDGIIPGNLKRKTAKENAADVNGVTVSSQAISM 450

DB 385 VOFLRNITANDGGAIIYLGESGELSADYDGIIPGNLKRKTAKENAADVNGVTVSSQAISM 444
 QY 451 GSGGKITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 510
 DB 445 GSGGKITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIIVFANGSSTL 504
 QY 511 YQNTIIEGRIVLREKAKLSVNSLSQSGSLYMEAGSTWDFVTPQPPQPPAAANQLITLS 570
 DB 505 YQNTIIEGRIVLREKAKLSVNSLSQSGSLYMEAGSTLDFVTPQPPQPPAAANQLITLS 564
 QY 571 NLHLSLSLLANNAVTPPTNPADSDHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDRYD 630
 DB 565 NLHLSLSLLANNAVTPPTNPADSDHPAIVIGSTTAGSVTISGPIPFEDLDDTAYDRYD 624
 QY 631 WLGSNKINVLKQLQGTKPPANAPSDLTGLNEMPKYQSGSKLAWDPNTANNPGYTLKA 690
 DB 625 WLGSNKINVLKQLQGTKPPANAPSDLTGLNEMPKYQSGSKLAWDPNTANNPGYTLKA 684
 QY 691 TWTGTGNGPGRVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNPFYH 750
 DB 685 TWTGTGNGPGRVASLVPNSLWGSILDIRSAHSAIOASVDGRSYCRGLWVSGVSNPFYH 744
 QY 751 DRDALGOGYRISGYSGLGANSYFGSSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS 810
 DB 745 DRDALGOGYRISGYSGLGANSYFGSSMFGLAFTVEFGRSKDYVVCRSNHHACIGSVYLS 804
 QY 811 TQOALCSYLPFGDAFIRASYFGNQHMKTSYTPAESDVRMDNCLAGEIGAGLPVITTP 870
 DB 805 TQOALCSYLPFGDAFIRASYFGNQHMKTSYTPAESDVRMDNCLAGEIGAGLPVITTP 864
 QY 871 SKLYNLRLPFPVQAEFSAADHESFTERGDOARAPKSGHLLNLSPVGVGKPDRCSTHPNK 930
 DB 865 SKLYNLRLPFPVQAEFSAADHESFTERGDOARAPKSGHLLNLSPVGVGKPDRCSTHPNK 924
 QY 931 YSFMAAVIDAYRTISGTETTLTSHQETWTTDAPHLARHGTVVGRGSMYASLTSNIEVYGH 990
 DB 925 YSFMAAVIDAYRTISGTETTLTSHQETWTTDAPHLARHGTVVGRGSMYASLTSNIEVYGH 984
 QY 991 GRYEYRDASRGYGLSAGSRVRF 1012
 DB 985 GRYEYRDASRGYGLSAGSKVRF 1006

RESULT 30
 ADD42756
 ID ADD42756 standard; protein; 670 AA.
 XX AC ADD42756;
 XX DT 15-JAN-2004 (first entry)
 XX Chlamydia pmpG passenger domain protein SEQ ID NO:169.
 DB Chlamydia infection; Chlamydia; antibiotic; antiinflammatory;
 KW antifertility; cardant; antiarteriosclerotic; ophthalmological;
 KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
 KW tubal obstruction; infertility; male infertility; ocular infection;
 KW blindness; acute respiratory tract infection; atherosclerosis;
 KW coronary heart disease.
 XX Chlamydia sp.
 OS WO2003041560-A2.
 PN 22-MAY-2003.
 XX 05-NOV-2002; 2002WO-US035624.
 PF 06-NOV-2001; 2001US-00012256.
 PR 05-DEC-2001; 2001US-00007693.
 PR 15-JUL-2002; 2002US-00197220.
 XX

QY 142 PAATNKGSQPTTTSTPSNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQG 201
 Db 63 PAATNKGSQPTTTSTPSNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQG 122
 QY 202 ISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSS 261
 Db 123 ISKLCVFOENTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSS 182
 QY 262 STSTEDPVVFSRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAKQ 321
 Db 183 STSTEDPVVFSRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAKQ 242
 QY 322 PTSQASNTSNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAK 381
 Db 243 PTSQASNTSNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAK 302
 QY 382 KLSVANCGPVFLNRIANDGAIYLGESGELSLADYGDIIYFDGNLKRKTAKENADVNGV 441
 Db 303 KLSVANCGPVFLNRIANDGAIYLGESGELSLADYGDIIYFDGNLKRKTAKENADVNGV 362
 QY 442 TVSSQASISMGSGGKIITLRAKAGHQILFNDPIEMANGNQPQSSKLLKINDGEGYTGDI 501
 Db 363 TVSSQASISMGSGGKIITLRAKAGHQILFNDPIEMANGNQPQSSKLLKINDGEGYTGDI 422
 QY 502 VPANGSSTLYQNVITBOGRIVLREKAKLSVNSLSQTGSLYMBAGSTWDFVTPQPQPP 561
 Db 423 VPANGSSTLYQNVITBOGRIVLREKAKLSVNSLSQTGSLYMBAGSTWDFVTPQPQPP 482
 QY 562 AANQLITLSNLHLSLLANNAVTPNTPNPPAQDS 597
 Db 483 AANQLITLSNLHLSLLANNAVTPNTPNPPAQDS 518

RESULT 32

AA16739
 ID AA16739 standard; protein; 505 AA.
 XX
 AC AA16739;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE C. trachomatis L2 HMW protein fragment.
 XX
 KW Chlamydia; high molecular weight protein; HMW protein; urethritis;
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
 KW salpingitis; tubal occlusion; infertility; cervical cancer;
 KW arteriosclerosis; atherosclerosis.
 XX
 OS Chlamydia trachomatis.
 XX
 FN WO9917741-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 01-OCT-1998; 98MO-US020737.
 XX
 PR 02-OCT-1997; 97US-00942596.
 XX
 PA (ANTE-) ANTEX BIOLOGICS INC.
 XX
 PI Jackson JW, Pace JL;
 XX WPI, 1999-287659/24.
 DR
 XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
 PT cervical cancer.
 XX
 XX Claim 5; Page 123-125; 141pp; English.
 XX
 CC The invention relates to an isolated Chlamydia species high molecular
 CC weight (HMW) protein having an apparent mol. wt. of 105-115 kD as

CC determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
 CC be used for preventing, treating or ameliorating a disorder related to
 CC Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
 CC lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
 CC pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
 CC cancer, infertility, arteriosclerosis and atherosclerosis. The products
 CC can also be used for detection and diagnosis. The present sequence
 CC represents a C. trachomatis HMW protein fragment
 XX
 SQ Sequence 505 AA;
 Query Match 48.4%; Score 2547; DB 2; Length 505;
 Best Local Similarity 99.8%; Pred. No. 2.4e-167;
 Matches 504; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 29 EIMVPOGIYDGETLTVSPFTYVIGDPGSGTTFVFSAGELTLKLNLSIAALPLSCFNLGSS 88
 Db 1 EIMVPOGIYDGETLTVSPFTYVIGDPGSGTTFVFSAGELTLKLNLSIAALPLSCFNLGSS 60
 QY 89 FTVLGRGHSITFENIRTSNKAALNSAADGLFTIEGKELSPNCNLSLAVLPAATTNK 148
 Db 61 FTVLGRGHSITFENIRTSNKAALNSAADGLFTIEGKELSPNCNLSLAVLPAATTNK 120
 QY 149 GSQPTTTTSPNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQGISKLCVF 208
 Db 121 GSQPTTTTSPNGTIYSKTDLLLNNEKFSFYSLVSGDGAIDAKSLTVQGISKLCVF 180
 QY 209 QENTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDP 268
 Db 181 QENTAQADGGACQVVTSPSAMANEAPAFVANVAGVGGGIAAVQDQGGVSSSTSTEDP 240
 QY 269 VVSFRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAKQPTSGQAS 328
 Db 241 VVSFRNTAVEFDGNVAVGGIYSGNVAFNNKTLFLNNVASPVYIAAKQPTSGQAS 300
 QY 329 NTSNNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAKLSVANC 388
 Db 301 NTSNNYGDGGAIFCKNGAQAGSNNGSVSPDGEVVPFSSNVAAGKGAIYAKLSVANC 360
 QY 389 GPVQFLNRIANDGAIYLGESGELSLADYGDIIYFDGNLKRKTAKENADVNGVTVSSQAI 448
 Db 361 GPVQFLNRIANDGAIYLGESGELSLADYGDIIYFDGNLKRKTAKENADVNGVTVSSQAI 420
 QY 449 SMGSGGKIITLRAKAGHQILFNDPIEMANGNQPQSSKLLKINDGEGYTGDIIVFANGSS 508
 Db 421 SMGSGGKIITLRAKAGHQILFNDPIEMANGNQPQSSKLLKINDGEGYTGDIIVFANGSS 480
 QY 509 TLQNVTIEQGRIVLREKAKLSVNS 533
 Db 481 TLQNVTIEQGRIVLREKAKLSVNS 505
 RESULT 33
 AA16751
 ID AA16751 standard; protein; 458 AA.
 XX
 AC AA16751;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE Chlamydia HMW protein fragment.
 XX
 KW Chlamydia; high molecular weight protein; HMW protein; urethritis;
 KW bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
 KW cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
 KW salpingitis; tubal occlusion; infertility; cervical cancer;
 KW arteriosclerosis; atherosclerosis.
 XX
 OS Chlamydia sp.
 XX
 FN WO9917741-A1.
 XX
 PD 15-APR-1999.

```

XX 01-OCT-1998; 98WO-US020737.
XX
XX 02-OCT-1997; 97US-00942596.
XX
XX (ANTE-) ANTEX BIOLOGICS INC.
XX
XX Jackson JW, Pace JL;
XX
XX WPI; 1999-287659/24.
XX
XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
XX cervical cancer.
XX
XX Claim 5; Page 133-135; 141pp; English.
XX
XX The invention relates to an isolated Chlamydia species high molecular
XX weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
XX determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
XX be used for preventing, treating or ameliorating a disorder related to
XX Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
XX lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
XX pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
XX cancer, infertility, arteriosclerosis and atherosclerosis. The products
XX can also be used for detection and diagnosis. Sequences AAY16740-Y16752
XX represent Chlamydia HMW protein fragments
XX
XX Sequence 458 AA;
XX
XX Query Match 44.6%; Score 2350; DB 2; Length 458;
XX Best Local Similarity 100.0%; Pred. No. 9e-154;
XX Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 217 GGACQVVTFSANANAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVVFSRNT 276
XX 1 GGACQVVTFSANANAPAFVANVAGVGGGIAAVQDGGQGVSSSTSTEDPVVFSRNT 60
XX
XX 277 AVEFDGNVAVGGGIYSGNVAFLNNGKTLFLNNVAVPVYIAKQPTSGOASNTSNYGD 336
XX 61 AVEFDGNVAVGGGIYSGNVAFLNNGKTLFLNNVAVPVYIAKQPTSGOASNTSNYGD 120
XX
XX 337 GGAIFCKNGAQAGSNNSGVSFDGEGVWPFSSNVAAGKGAIVAKKLSVANCGPVQFLRN 396
XX 121 GGAIFCKNGAQAGSNNSGVSFDGEGVWPFSSNVAAGKGAIVAKKLSVANCGPVQFLRN 180
XX
XX 397 IANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVNGVTVSSQAIISMGSGGKI 456
XX 181 IANDGGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVNGVTVSSQAIISMGSGGKI 240
XX
XX 457 TTLRAKAGHQLFNDPIEMANGNQPAAQSSKLLKINDGEGYTGDIVPANGSSTLYQNVTI 516
XX 241 TTLRAKAGHQLFNDPIEMANGNQPAAQSSKLLKINDGEGYTGDIVPANGSSTLYQNVTI 300
XX
XX 517 EQGRIVLREKAKLSVNSLSOTGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 576
XX 301 EQGRIVLREKAKLSVNSLSOTGSLYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSL 360
XX
XX 577 SSSLANNVATNPPTNPPAOSHFAVIGSTTAGSVTISGPIFFEDLDODTAYDRYDWLGSNQ 636
XX 361 SSSLANNVATNPPTNPPAOSHFAVIGSTTAGSVTISGPIFFEDLDODTAYDRYDWLGSNQ 420
XX
XX 637 KINVLKQLGTGKPPANAPSDLTIGNENPKYGYQGSWKL 674
XX 421 KINVLKQLGTGKPPANAPSDLTIGNENPKYGYQGSWKL 458
XX
XX RESULT 34
XX AAY16752
XX ID AAY16752 standard; protein; 325 AA.
XX
XX AC AAY16752;
XX
XX 21-JUL-1999 (first entry)

```

```

XX Chlamydia HMW protein fragment.
XX
XX Chlamydia; high molecular weight protein; HMW protein; urethritis;
XX bacterial infection; conjunctivitis; lymphogranuloma venereum; LGV;
XX cervicitis; epididymitis; endometritis; pelvic inflammatory disease; PID;
XX salpingitis; tubal occlusion; infertility; cervical cancer;
XX arteriosclerosis; atherosclerosis.
XX
XX Chlamydia sp.
XX
XX WO9917741-A1.
XX
XX 15-APR-1999.
XX
XX 01-OCT-1998; 98WO-US020737.
XX
XX 02-OCT-1997; 97US-00942596.
XX
XX (ANTE-) ANTEX BIOLOGICS INC.
XX
XX Jackson JW, Pace JL;
XX
XX WPI; 1999-287659/24.
XX
XX New Chlamydia protein useful for treating conjunctivitis, urethritis and
XX cervical cancer.
XX
XX Claim 5; Page 135-136; 141pp; English.
XX
XX The invention relates to an isolated Chlamydia species high molecular
XX weight (HMW) protein having an apparent mol. wt. of 105-115 kD as
XX determined by SDS-PAGE. The HMW proteins and nucleic acid molecules can
XX be used for preventing, treating or ameliorating a disorder related to
XX Chlamydia e.g. bacterial infection, conjunctivitis, urethritis,
XX lymphogranuloma venereum (LGV), cervicitis, epididymitis, endometritis,
XX pelvic inflammatory disease (PID), salpingitis, tubal occlusion, cervical
XX cancer, infertility, arteriosclerosis and atherosclerosis. The products
XX can also be used for detection and diagnosis. Sequences AAY16740-Y16752
XX represent Chlamydia HMW protein fragments
XX
XX Sequence 325 AA;
XX
XX Query Match 32.9%; Score 1735; DB 2; Length 325;
XX Best Local Similarity 100.0%; Pred. No. 2e-111;
XX Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 688 LKATWTKTGYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSVNF 747
XX 1 LKATWTKTGYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGSVNF 60
XX
XX 748 FTHDRDALCGGYRIYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRSNHHACIGSV 807
XX 61 FTHDRDALCGGYRIYISGGYSLGANSYFGSSMFGLAFTVFGRSKDYVVCRSNHHACIGSV 120
XX
XX 808 YLSTQALCGSYLFGDAPIRASVFGNQHMKTSTYTPAESDVWRDNNCLAGEIGAGLPIV 867
XX 121 YLSTQALCGSYLFGDAPIRASVFGNQHMKTSTYTPAESDVWRDNNCLAGEIGAGLPIV 180
XX
XX 868 ITPSKLYLNLRLPFVQAQFSAADHESFTEEGDQARAFKSGHLLNLSPVGVCKFDRCSSTH 927
XX 181 ITPSKLYLNLRLPFVQAQFSAADHESFTEEGDQARAFKSGHLLNLSPVGVCKFDRCSSTH 240
XX
XX 928 PNKYSFMAAYICDAYTISGTETLLSHQETTTTDAFLARHGTVVGRGSMYASLTNSIEV 987
XX 241 PNKYSFMAAYICDAYTISGTETLLSHQETTTTDAFLARHGTVVGRGSMYASLTNSIEV 300
XX
XX 988 YGHGRYEYRDASRGYGLSAGSRVRF 1012
XX 301 YGHGRYEYRDASRGYGLSAGSRVRF 325
XX
XX RESULT 35

```


XX WPI; 1999-105610/09.
 DR N-PSDB; AAX06821.
 XX
 PT Species-specific test for identifying mammals infected with Chlamydia
 PT pneumoniae - comprises detecting antibodies specific for outer membrane
 PT proteins of C. pneumoniae or nucleic acids encoding these proteins.
 XX
 PS Claim 7; Page 56-58; 115pp; English.
 XX
 CC This polypeptide comprises the novel 96.7 kDa surface exposed protein
 CC Omp9 of the human respiratory pathogen Chlamydia pneumoniae. Its amino
 CC acid sequence was deduced from DNA (see AAX06821) isolated from a C.
 CC pneumoniae expression library. The invention provides 12 novel surface
 CC exposed proteins, Omp4-Omp15 (see AAX06821-27). A new species specific test is
 CC sequences encoding them (see AAX06821-27). The test comprises detecting antibodies specific
 CC Chlamydia pneumoniae. The test comprises detecting antibodies specific
 CC for Omp4-Omp15 or detecting nucleic acid fragments encoding these outer
 CC membrane proteins, especially by PCR. The proteins are also used in the
 CC diagnosis of C. pneumoniae infection in mammals. The nucleic acids and
 CC proteins can also be used in the immunization of mammals, the nucleic
 CC acids being particularly useful as DNA vaccines for effecting in vivo
 CC expression of antigens. The vaccines may also prevent atherosclerosis and
 CC bronchial asthma, which are possibly associated with C. pneumoniae.
 CC (updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 918 AA;
 Query Match 21.7%; Score 1143; DB 2; Length 918;
 Best Local Similarity 30.9%; Pred. No. 9.1e-70;
 Matches 324; Conservative 163; Mismatches 393; Indels 170; Gaps 34;
 QY 1 MOTSFHKPFLSMILAYSCSLNGGGYAAEIMVPPQGIYDGETLTVSP-PYTVIGDSGGTVV 59
 DB 1 MRSSFLLISSLAPFLMVSADAADLTGSRDSYNGDTSTTEPKAATSDASGTTY 60
 QY 60 FSAGELTKNLNLSIAALPLCPGNLLGSFTVLGRHSILTFENIRST-NGAALNSRAD 118
 DB 61 ILDGVSISQAGKQ-TSLTTSCTNTAGNLTLFNGFSLFHPDNISSIVAGVVSNTAAS 119
 QY 119 GLPTIEGKELSPNCNLSLAVLP-AATNKGSTPTTTPSPNGSIYKTDLLILNNEK 177
 DB 120 GITKPSGF-----STRLMAAPRTTGKGAIKT-----DGLVPSIGNLDQN--- 161
 QY 178 FSPYSLNVLGSGDAIDAKSLITVQGISKLCPVQNTAQADGACQVVTFSAMANEAPIAF 237
 DB 162 -----ENASSENAGAIKTLTSLGSTRF-----VAF 188
 QY 238 VANVAGVRGGIAAODGQGVSSSTSTEDPVVPSRNTAVEPDGNVARVGGIYSGNV 297
 DB 189 LGNSSSQGGAIYASGD-----SVISENAGILSFGNNSA-----TTSGGAIASGNL 235
 QY 298 AFLNKGKTLPLNNVSPVYIAAQPTISGQASNTSNYGDGAIKCKGAGQAGNNSGSVS 357
 DB 236 VTISNNQIFP-----DGCKATT-----NGGAIKCN-----KAGANPPDILT 271
 QY 358 FDCEGVVFFSSNVAAGKGAIAKLSVAN-CGPVQFLRNIAAND-----CGAIYLGESGEL 412
 DB 272 LSGNESLHFLNNTAGNSGGAIYTKLVLSGRGVLFSNNKAANATPKGGAIALDSGEI 331
 QY 413 SLSADYGDIIIFDGNLKRKTAKENADVNGVTSSQASIMSGGKITTIRAKAGHQILNDP 472
 DB 332 SISADLGNIIIFEGNTTSTTSGSPA-----SVTRNAIDLASNAKFLNLRGKNKVIYDP 385
 QY 473 IEWANGNNQPAQSKLLKINDGEG--YTGDIVA-----NGSSTLYQNVITIE 517
 DB 386 ITSSGATDK-----LSLNKADAGSGNTYEGYIVFSGEKLSEELKPKDNLKSTTQAVELA 441
 QY 518 QGRVILREKAKLSVNSLSQTGGG-LYNEAGTSWDFVTPPPQPPAPPAQLITLSNLHSL 576
 DB 442 AGALVLKDGVTVVNTITQEGSKVMDGGTTFE-----ASAEGLTGLNLAINI 490

QY 577 SSLLANNVNTNPTTNPAPQDSHPAVIGSTTAG-SVTISGPIFFEDLDDTADVRDVLGSLN 635
 DB 491 DSLGDTN-----KAIIRATAASKDVALSGPIMLVDAQGNYYEHHN-LSQQ 534
 QY 636 QKINVLKQL-GTKEPPANAPSDLTGLNEMPKYQGSWKLAW-DENTANNNGPYTLKATWT 693
 DB 535 QVFFLIELSAQGTMTTDDIP-DTPIINTNTNHYGQGTGIIIVVDDATAKTAKNATL--TWT 591
 QY 694 KTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDR 753
 DB 592 KTGKXPERQGPLVPNSLWGSFVDVRSIQSLMDRSTSSLSSTNLWVSGIADFLHEDQK 651
 QY 754 ALGOQYRISYGGISLGANSYFGS-SMGLAPTEVPGRSKDYVCKSNHHACIGSV---YL 809
 DB 652 QNQRSHRSAGYALGGGFFTAENFFNFAFCQLFGYDKDLVAKNTHVTVAGAMSRYRL 711
 QY 810 STQOALCQSYLFQDA----FI---RASVGFQNHMKTSYTPAESDVWDNNCLAGEIGA 862
 DB 712 GESKTL-AKILSGNSDSLFPFVNARFAYGHTDNNMTTKYGVSPVKSGWGNDAFGIECG 770
 QY 863 GLPIVITPSKLVNLRELRFVQAEFSYADHESPTERGDQARAFKSGHLLNLSVPVGVKFD 922
 DB 771 AIPVVASGRGWVDHTFPLEMIYAHQNDKENGTEGRSFQSEDLFNLAVPGIKPEK 830
 QY 923 CSSTHPNKYSFMAAYICDAYRTISCTETTLTSHQETWTTDAFHLARHGTVVVRGSMYASLT 982
 DB 831 PSD--KSTYDLSIAVVPDVRNDPGCTTTLMVSGDSWSTCGTSLSRQALLVRAGNHAPA 888
 QY 983 SNIIEVGHGRVYRDASRGYGLSAGSRVRF 1012
 DB 889 SNFEVFSQFEVELGRSSRYAIDLGRFGF 918
 RESULT 39
 AAY94327
 ID AAY94327 standard; protein; 928 AA.
 XX
 AC AAY94327;
 DT 12-SEP-2003 (revised)
 DT 11-AUG-2000 (first entry)
 XX Chlamydia pneumoniae 98kD putative outer membrane protein.
 XX Chlamydia; antigen; vaccine; infection; outer membrane protein.
 XX Chlamydia pneumoniae.
 XX WO200026237-A2.
 XX 11-MAY-2000.
 XX 29-OCT-1999; 99WO-GB003579.
 XX 29-OCT-1998; 98US-0106070P.
 XX 01-MAR-1999; 99US-0122066P.
 XX 27-OCT-1999; 99US-00428122.
 XX (CONN-) CONNAUGHT LAB LTD.
 XX Murdin AD, Oomen RP, Dunn PL;
 XX WPI; 2000-365569/31.
 DR N-PSDB; AAA27021.
 XX Novel Chlamydia 98 kDa putative outer membrane protein antigen, used for
 XX vaccination and protection against Chlamydia infection.
 XX Claim 6; Fig 1; 93pp; English.
 XX The present sequence is the 98kDa putative outer membrane protein from
 CC Chlamydia pneumoniae. The genomic sequence was amplified using two PCR
 CC primers. The 5' primer contains a NotI restriction site, a ribosome

CC binding site, an initiation codon and a sequence close to the 5' end of
 CC the 98kDa putative outer membrane protein coding sequence. The 3' primer
 CC contains the sequence encoding the C-terminal sequence of the putative
 CC outer membrane protein and a BglI restriction site. The stop codon was
 CC excluded and an additional nucleotide was inserted to obtain an in-frame
 CC C-terminal fusion with the Histidine tag. The PCR product was cloned into
 CC a eukaryotic expression vector (pCA-Myc-His) by restricting both the
 CC vector and the PCR product with NotI and BamHI and performing a ligation
 CC reaction. This expression vector was injected intramuscularly and
 CC intranasally into mice, which were subsequently inoculated with Chlamydia
 CC pneumoniae. The chlamydial lung titers of the immunised mice were lower
 CC than those of the controls. Thus the 98kDa putative outer membrane
 CC protein can be used as a vaccine to provide protection against Chlamydia
 CC infections, especially Chlamydia pneumoniae infections. The present
 CC polypeptide may also be administered orally to treat Chlamydia infection.
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 928 AA;

Query Match 21.4%; Score 1128; DB 3; Length 928;
 Best Local Similarity 29.2%; Pred. No. 1e-68;
 Matches 311; Conservative 169; Mismatches 394; Indels 192; Gaps 32;

QY 1 MOTSPKFFLSMILAYSCSLNGGGYAAE-IMVPGIYDGETLTVSFPVTVIGDPSGTVV 59
 DB 1 MKSSFPKVFSTFAIPPLSMI-----ATETVLDSASFQGNK-NGNFSVRSQEDAGITY 54
 QY 60 FSAGELTLKLNLDNSIAALPLSCFGLNLSFTVLGRHSLTFPENIRTS-NGAALSNSAAD 118
 DB 55 LPKGNVTLENIPTGTATKSCFNNTKGDLTPTGNGSLLFTVDAGTVAGAAVNSSVVD 114
 QY 119 GLFTIEGKELSFNSCNLLAVLPAATNKGSTPTTSTPNSGTYISKTDLLLNNEKF 178
 DB 115 KSTTFIGFSSLSF-----IASPSSITTKGAVSCSTGS-----LSLTKNYSL 157
 QY 179 SPYBNLVSDGGAIDAKSLTVGIGKLCVFBQNTAQDGGACQVVTSPSAMANEAPIAV 238
 DB 158 LPKFNFTDNGAIYAKLSLTLGTTMSALFSENSSKGGALQTSDAUTYIGNQGEVSF- 216
 QY 239 ANVAGVGGGIAAVDGGQGVSSSTSTBPDVVSFSRNTAVEPDGNVARVGGIYSYGNVA 298
 DB 217 -----SDNTSSDS-----GAAIFTEASVT 235
 QY 299 FLNNGKTLFLNVAAPVYIAKQPSQASNTSNYDGGALFCKNGAQAQSNNGSVSP 358
 DB 236 ISNNAKVGFIDN-----KVTGASSSTTGDH-SGGAICAYK-----TSTDTKVTL 278
 QY 359 DQGVVFPSSNVAAGKGAIAKLSVANGCPVQFLRNIA-----DGAIVLGSSELS 413
 DB 279 TGNQMLLFSNNTSTTGAIAIVYKLELAGGLTLFSRNSVNGGTAPKGAIAIEDSGELS 338
 QY 414 LSADYGDIIIPGNLKRATKENAADVNGVTVSQAISMGSGGKIITLRAKAGHQLFNDDPI 473
 DB 339 L6ADSGDIVFLGN---TVSTSTPGTN-----RSSIDLGTSAKMTALRSAGRAIVFYDPI 390
 QY 474 EWANGNQPQASKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNTV 515
 DB 391 TTGSSTT---VTDVLKVNETHPADSALQYTGNIIFTGKLSGETEAADSKNLTSLQPV 446
 QY 516 TEQGRIVLREKAKLSVNSLSOTGGS-LYMEAGSTWDFVTTPQPPQPAANQLITLSNHL 574
 DB 447 LSGGTLSLKHGVTQTQFTQADSRLEMDVGTLL-----PADTS---TINNLI 494
 QY 575 SLSSILLANNAVTPNPPNPAQDSHPAVIGS-TTAGSVTISGTFPEDLDDTDAYDRYWL 633
 DB 495 NISSI-----DQAKKAIETKATSKNLTLSGTTITLLDPTGTFYENHS-LR 538
 QY 634 SNQKINVLKQL-GTKPPANAPSDLTGLNEMPKYQGSW-KLAWDPNTANNPPTLKAT 691
 DB 539 NPQSYDILELKASGVTSTAVTPDPIMGKEF-HYGYQGTWGPVWGTGASTTATP-----N 593
 QY 692 WTKTYNPPQPRVAVSLVPSNLWGSITLDIRSAHSAIQASVDGRSVCRLWVSGVSNFFVHD 751

DB 594 WTKTYNPPQPRVAVSLVPSNLWNAFIDISSILHYLMETANEGLQGRFACWAGLSNFFHKD 653
 QY 752 RDALQGGVRYISGGYSLGANSYFGS-SMFGLAFTFVGRSKDYVVCRSNHHACIGSVYLS 810
 DB 654 STKTRRGPHLSGGYVIGGNLHTCSKILSAAFQQLFGDRDRDYFAKNQ-----GTVYGG 708
 QY 811 TQQAALCGSYLPGDAFI-----PASVCPGNHMKTSYTFABE 846
 DB 709 TLY-----YQHNEYIYSLPCKRLPCSLSYVTEIPVLFSGNLSYTHTDNDLTKTYTPT 763
 QY 847 SDVRWNNCLAGEIGAGLPIVITPSKLYNLRLRPFVQAEPSYADHESFTEGDQARAFKS 906
 DB 764 VKGSWENDSPALEFGGRAPICLDESALF-EQYMPKQLQFYAHQEGFKQGTAREFGS 822
 QY 907 GHLNLNVVPVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETTLTLLSHQBTWTDAFHL 966
 DB 823 SRLVNALPIGIRFDKESDCQDATYNTLTGLYTVDLVRSNPDCVTTTLRIGSDSKWTKFTGNL 882
 QY 967 ARHGVTVRGSMYASLTSNIERYGHRVYRDRASRGYGLSAGSRVRF 1012
 DB 883 ARQALVLRAGNHFCFNSNFPAFSQFSELRGSSRNYNVDLGAKYQF 928

RESULT 40
 ABP56019 standard; protein; 926 AA.

AC ABP56019;
 DT 29-AUG-2003 (revised)
 DT 25-FEB-2003 (first entry)
 DE Chlamydia psittaci antigen CP4#12 protein SEQ ID NO:57.
 KW Chlamydia psittaci; vaccination; vaccine; antigen; immune response;
 KW immunisation; antibacterial; infection.
 OS Chlamydophila caviae.

PN WO200253588-A2.
 PD 11-JUL-2002.

PF 17-DEC-2001; 2001WO-US048715.

PR 15-DEC-2000; 2000US-00738269.

PR (TEXA) UNIV TEXAS SYSTEM.

PI Johnston SA, Stemke-Hale K, Sykes KF, Kaltenboeck B;

DR WPI; 2002-537942/57.

DR N-PSDB; ABQ84775.

PT Vaccine for immunization of animal, preferably bovine, against Chlamydia
 psittaci, comprises at least one polynucleotide having a C. psittaci
 sequence, or at least one C. psittaci antigen.

PS Claim 9; Page 158-161; 164pp; English.

CC The present invention describes a vaccine (I) for the immunisation of an
 animal against Chlamydia psittaci comprising at least one polynucleotide
 (Ia) having a C. psittaci sequence, or at least one C. psittaci antigen
 (Ib), and a carrier. (Ia) and (Ib) have antibacterial activity. (I) is
 useful for the immunisation of a bovine. The present sequence represents
 a C. psittaci antigen from the present invention. (Updated on 29-AUG-2003
 to standardise OS field)

XX Sequence 926 AA;

Query Match 21.3%; Score 1122; DB 5; Length 926;
 Best Local Similarity 29.7%; Pred. No. 2.6e-68;
 Matches 323; Conservative 152; Mismatches 376; Indels 236; Gaps 34;

QY 1006 AGSRVRF 1012
Db 920 LGAKVAF 926

Search completed: May 13, 2006, 12:15:00
Job time : 200 secs

QY 1 MCTSHKPLSLMILAYSCSLNGGVAABIMVPOGIYDGETILTVSPPTVI----- 51
Db 1 MRPSLYKILISTL-----TLPISFHSQLAHEVALTQE 34
QY 52 -----GDPGGTTFVSAGBELTKNLNDSI-----AALPLSCFNLGSLFTVLGRHS 97
Db 35 SILDANGAFPOSTAGG-TIYNVEDSIVDVGQTAALASSAFVQTADNLTFKGNHS 93
QY 98 LTFENIRTSNGAALSNAADGLFTIEGKELSFNSCNSLLAVLPAATYTKGSQTPPTS 157
Db 94 LSIITNAGANPAGINVTADKILTLTDFSKLSFKEC-----PSSLVNTG----- 138
QY 158 TFSNGTIYSKTDILLILNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOENTAQADG 217
Db 139 -----KGAKSGGALNLANNASILFDQNYSAENGAIISCKAFSLTGSSEISFTTNSTAKG 195
QY 218 GACQVVTFSAMANEAPIAFVANVAGVGGIAAVODGQGVSSSTSTEDPVVPSRNTA 277
Db 196 GA-----IAATGIAHLSNQ-----GTIRFSGNTA 220
QY 278 VEPDGNVAVGGIYSGNVAFNLNGKTLFLNNVASPVVIAAKQPTSGOASNTSNYGDG 337
Db 221 V-----NSGGAVYSEASMTIAGNNHVAFSNNAVS-----GSSDGGC 256
QY 338 GAIFCKNGAQAAGNSGVSVPDGEVVPFSSNVAAAGGGAIIYAKKLSVANCGPVQFLRN- 396
Db 257 GAHCSK-----TGSAPTILTIIRDKNVLIPEENTSSAKGGAITYDKLILTSGGGTAFINNK 311
QY 397 ---IANDGGAIVLGBESGELSADYDGIIPDGNLKRKTAKENAADVNGVTVSSQASISMSG 453
Db 312 VTHATPKGAIGIANGECSLTAHGDITFDNNLMAT-QDNA-----TIKRNAINIEGN 364
QY 454 GKITTILRAKAGHOILFNDPIEMANGNNQPAQSKLLKINDGEG---YTGDIVPA----- 504
Db 365 GKFPVNLRAASGKTSIFVDPIV-EGN-----AADLLTLNKAEGDKTYNGRIIFSGEKLTE 418
QY 505 -----NGSSTLYQNVTTIEQRIVLREKAKLSVNSLSQSGSL-YMEAGSTWDFVTPQP 556
Db 419 EQAAVADNLKTTPTPITLAAGBLVLRSGVEBAKTVVQTAGSLILMDAGT----- 469
QY 557 PQQPPAANQLITLSNLHLSLSSLANNAVTPNPPTNPPAODSHP-AVIGSTTAG-SVTISG 614
Db 470 -----KLSAKTEDATLTNLAINENTLDGKFAVVDVAAGKNVTLSG 511
QY 615 PIFFDLDDTAYDRYDWLGSNOKIN-----VLKLQ-----GTKPPANAPSDLTGNEMPKYG 667
Db 512 AIGVIDPTGKFYE-----NHKLNLTALGGIQLSGKSGSVTTTNNVPSHV-VGVAETHYG 563
QY 668 YQGSWKLAWDPNTANNOPYTLKA---TWTKTYNPGPERVASIYPNSLWGSILDIRSAHSA 725
Db 564 YQGNWSVSWYKDN-NSDPKTQTAIFTWNKTGYVPNPERRAPLVLNLSWGSFIDLRISQDV 622
QY 726 IQASVDG-RSYCRGLWSGVSNPFYHDDALGQGYRISGYSILGANSYFG-SSMFGIAP 783
Db 623 LERSVDSILETRGLWVSGIGNPFHKORNAENKFRHISSGYVLGATNTSREDSLSVAF 682
QY 784 TEVFRSKDYVVKRSHHACITGSYV-----LSTQQALCGSY-----LFGDAF 825
Db 683 COLPAKDKDYLVSNAANVYAGSVYQHVSKFDDLTFLFNGPNTCCSGFSKEIPIPLDAQ 742
QY 826 IRASYFGNQHMTSYTFAESDVVRWNNCLAGIAGLPIVITPYSKLYLNELPFVQAE 885
Db 743 I--TYCHTANNMTTSYTDYEVKSGWGNLTGLTLSTSVPIPVFSSSIP-DSYAPFAKLQ 799
QY 886 PSYADHESPTTEGQOARAFKSGHLLNLVSPVGVKFDRCSTHPPNKYSFMAAYICDAYRTI 945
Db 800 VVYAHQDDFKBPTTEGRVFESSDDLNVSPVIGIKPEKLSYGERSAIDLTLMYIPDVYRHN 859
QY 946 SGTETLLSHOETWTTDAFLHARGVVRGSMYASLTSNIEVYGHGREYRDASRGYGLS 1005
Db 860 PSCMTGLAINDVSWLTATNLARQAFIVRAGNHIALTSVEMFSQFGFELRSSRNTNVD 919

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:17:59 ; Search time 47 Seconds
(without alignments)
2071.731 Million cell updates/sec

Title: US-10-701-844-2
Perfect score: 5267
Sequence: 1 MQTSFHKFPLSNILAYSCS.....VEYRDASRGYGLSAGSRVRF 1012

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5131.5	97.4	1013	2 G71460	probable outer mem
2	3845	73.0	987	2 H81722	polymorphic membra
3	1365.5	25.9	973	2 B86547	polymorphic outer
4	1365.5	25.9	973	2 P72076	polymorphic outer
5	1365.5	25.9	935	2 C81593	polymorphic membra
6	1101	20.9	928	2 G86546	polymorphic outer
7	1101	20.9	928	2 G81591	polymorphic membra
8	1090	20.7	928	2 D72077	polymorphic outer
9	1090	20.7	928	2 H86546	polymorphic membra
10	1090	20.7	949	2 P81591	polymorphic outer
11	1083	20.6	930	2 D86546	polymorphic membra
12	1083	20.6	930	2 A81591	polymorphic outer
13	1079	20.5	930	2 D72078	polymorphic membra
14	1041	19.8	936	2 C72078	polymorphic outer
15	1040	19.7	936	2 B81591	polymorphic membra
16	1040	19.7	936	2 C86546	polymorphic outer
17	1021	19.4	928	2 B72077	polymorphic membra
18	1021	19.4	928	2 B86546	polymorphic outer
19	1012	19.2	1276	2 B86546	polymorphic membra
20	1012	19.2	1276	2 C81591	polymorphic outer
21	985.5	18.7	1407	2 B72078	polymorphic membra
22	978.5	18.6	772	2 H86492	Pmp_3 [imported] -
23	967	18.4	922	2 B72131	polymorphic outer
24	967	18.4	922	2 B86491	polymorphic membra
25	966	18.3	922	2 F81539	polymorphic outer
26	939.5	17.8	841	2 B72130	polymorphic membra
27	789	15.0	867	2 F81721	polymorphic membra
28	771.5	14.6	878	2 B71460	probable outer mem
29	703	13.3	712	2 B86492	polymorphic outer

ALIGNMENTS

RESULT 1

G71460
probable outer membrane protein G - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: G71460
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: G71460
A:Molecule type: DNA
A:Residues: 1-1013 <ARN>
A:Cross-references: UNIPROT:O84879; UNIPARC:UPI0000131CF5; GB:AE001360; GB:AE001273; NIT A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: pmpG

Query Match	97.4%	Score	5131.5	DB 2	Length	1013			
Best Local Similarity	97.0%	Pred. No.	8.6e-304						
Matches	983	Conservative	18	Mismatches	11	Indels	1	Gaps	1

QY	1	MQTSFHKFPLSNILAYSCCSLNGGVAIRIMVPOGIYDGETLTVSPYTVIGDPGSGTVTF	60
DB	1	MQTSFHKFPLSNILAYSCCSLNGGVAIRIMVPOGIYDGETLTVSPYTVIGDPGSGTVTF	60
QY	61	SAGELTLKNLDSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNCAALSNASDGL	120
DB	61	SAGELTLKNLDSIAALPLSCFNLGSLFTVLGRGHSITFENIRTSNCAALSDSANSGL	120
QY	121	FTIEGPKELSFNCNSLLAVLPAATNKGSGQPTTSTPSNGTIYSKTDLLILNNEKPSF	180
DB	121	FTIEGPKELSFNCNSLLAVLPAATNKGSGQPTTSTPSNGTIYSKTDLLILNNEKPSF	180
QY	181	YSLNLSGDDGADAKSLTVQGIKLCVCFQNTAQADGGACQVVTFSAMANEAPIAFVAN	240
DB	181	YSLNLSGDDGADAKSLTVQGIKLCVCFQNTAQADGGACQVVTFSAMANEAPIAFIAN	240
QY	241	VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVSFRNTAVBFDGNVARVGGGIYSYGNVAF	300
DB	241	VAGVRGGGIAAVQDGGQGVSSSTSTEDPVVSFRNTAVBFDGNVARVGGGIYSYGNVAF	300
QY	301	NGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGYGGGAIFCKNGAQ-AGSNNSGSVSP	359
DB	301	NGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAAGSNNSGSVSP	360
QY	360	GEQVWFSSNVAAGKGAIAKLSVANGCPVOFLNIANDGGAIVLGGSELGLSADYG	419
DB	361	GEQVWFSSNVAAGKGAIAKLSVANGCPVQFLNIANDGGAIVLGGSELGLSADYG	420
QY	420	DIIFDGNLKRKATENAADVNGVTVSQAISWGGGKITTLRAKAGHQIILFNDPIEMANG	479

```
Db 421 DIIPGNLKRATAENADVGVTVSSQALSMGSGKITTILRAKAGHQILFNDPIEMANGN 480
Qy 480 NQPAQSSKLLKINDGEGTGDIVFANGSGSTLYQNTVIEQGRIVLRBEKAKLSVNSLSQTGG 539
Db 481 NQPAQSSBPLKINDGEGTGDIVFANGSGSTLYQNTVIEQGRIVLRBEKAKLSVNSLSQTGG 540
Qy 540 SLYMEAGSTWDPVTPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 599
Db 541 SLYMEAGSTWDPVTPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 600
Qy 600 AVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSGNQKINVLKQLGQTKPPANAPSDTL 659
Db 601 ALIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSGNQKIDVLKQLGQTPSANAPSDTL 660
Qy 660 GNEPMKYGYQGSWKLAMPNTANNQPYTLKATWTKTGNPGRVAVLVPNSLWMSIIDI 719
Db 661 GNEPMKYGYQGSWKLAMPNTANNQPYTLKATWTKTGNPGRVAVLVPNSLWMSIIDI 720
Qy 720 RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDLALGQGYRIISGGYSLGANSYFGSSMF 779
Db 721 RSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDLALGQGYRIISGGYSLGANSYFGSSMF 780
Qy 780 GLAPTEVFGRSKDYVYCRSNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 839
Db 781 GLAPTEVFGRSKDYVYCRSNHACIGSVYLSQALCGSYLFGDAFIRASYGFGNQHMKT 840
Qy 840 SYTPABESDVRWNNCLAGEIGAGLPVITPSKLYNELRPFVQABPYADHESPTERGD 899
Db 841 SYTPABESDVRWNNCLAGEIGAGLPVITPSKLYNELRPFVQABPYADHESPTERGD 900
Qy 900 QARAPKSHLANLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETW 959
Db 901 QARAPKSHLANLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETW 960
Qy 960 TTDAFLARHGVMVGRSGVSYASITSNIEYVGHGRYEYRDSRGYGLSAGSRVRF 1012
Db 961 TTDAFLARHGVMVGRSGVSYASITSNIEYVGHGRYEYRDSRGYGLSAGSRVRF 1013

RESULT 2
H81722
polymorphic membrane protein G family TC0263 [imported] - Chlamydia muridarum (strain N4
C/Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: H81722
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: AB1500; MUID:20150255; PMID:10684935
A/Accession: H81722
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-987 <TET>
A/Cross-references: UNIPROT:Q9PL45; UNIPARC:UPI0000057867; GB:AE002293; GB:AE002160; NID
A/Experimental source: strain N199 (MoPn)
C/Genetics:
A/Gene: TC0263

Query Match 73.0%; Score 3845; DB 2; Length 987;
Best Local Similarity 72.6%; Pred. No. 1.3e-225;
Matches 739; Conservative 105; Mismatches 136; Indels 38; Gaps 13;
Qy 1 MQTSPHFKPFLSMILAYSCSLGGGYAAIBVWPGIYDGETLTVSPFTYVIGDPSGTTVP 60
Db 2 MQTSPHFKPFLAMLSY---SLQGGHADIEMPPGCIYDGTTLTAPPTVTVIGDPRGTVT 58
Qy 61 SAGELTLKNLDSIALPLSCFNLGSGFTVLGRHSLTPENIRTSNNGAALSNAADGL 120
Db 59 SSGLELKNLDSIATLPLSCFNLGSGFTVLGRHSLTPENIRTSNNGAALSNAADGL 118
Qy 121 FTIEGPKELSPSCNCSLLAVLPAATTKNGSQT-PTTSTSPNGTIYSNTDILLNNKFS 179
```

```
Db 119 FVIEAFBELSLNCSLSVSVF-----QTGTTTTVPNGHIYRTDILVLRDIKKVS 170
Qy 180 FYSNLVSGDGAIDAKSLTVQGISKLCPQFQENTAAQDGGACQVVTFSFSAANEAPFAVA 239
Db 171 FYSNLVSGDGAIDAKSLTVQGISKLCPQFQENTAAQDGGACQVVTFSFSAANEAPFAVA 230
Qy 240 NVAGVRCGGIAAVDQGGVSSSTSTEDPVVSPFRNTAVZPDGNVAVRGGIYSYGNVAP 299
Db 231 NVAGVRCGGIAAVDQGGVSSSTSTEDPVVSPFRNTAVZPDGNVAVRGGIYSYGNVAP 286
Qy 300 LANNKTLPLNNVAVSVTV--AAQPTSGQASNTSNNTGDGGAIFCKGGAQAGSNNSGSVS 357
Db 287 LGNAKTVFLSNVAVSVTVDPAA---ACGQPPADKNTGDGGAIFCKN-----DNTIGSVS 338
Qy 358 PDGEGVVPFSSNVAAKGGAIYAKKLSVANGCPVQFLRNTANDGGAIIYLGSGELSLAD 417
Db 339 PKDEGVVPFSSKNIYAAGKGGAIYAKKLTISDCGPVQFLGNVANDGGAIIYVQGBELSLAD 398
Qy 418 YGDIIFPQNLKRTAKENAADVGVTVSSQALSMGSGKITTILRAKAGHQILFNDPIEMAN 477
Db 399 RGDIIIPQNLKRMATQGAATVHDVWVASNAISMA7GQITTLRAKEGRRIILFNDPIEMAN 458
Qy 478 GNNQPAQSSKLLKINDGEGTGDIVFANGSGSTLYQNTVIEQGRIVLRBEKAKLSVNSLSQT 537
Db 459 G--QPV--IQTLTVNEGEGTGDIVFANGDNVLYSSIELSGQRIILREQTKLVNSLTQT 514
Qy 538 GGSLYMEAGSTWDPVTPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 597
Db 515 GGSVMEGGSTLDPVAV---TTPPAANSM-ALTWVHFSLASLLKNNGVTPPTNPVQVS 569
Qy 598 HPAVIGSTTAGSVTISGPIPFEDLDDTAYDRYDMLGSGNQKINVLKQLGQTKPPANAPSDL 657
Db 570 SPAVIGNTAAGTVTISGPIPFEDLDDTAYDNNQWLGADQITDVLQLHLGANPPANAPTDL 629
Qy 658 TLGNEMPKYGYQGSWKLAMPNTAN--NGPYTLKATWTKTGNPGRVAVLVPNSLW 714
Db 630 TLGNEMPKYGYQGSWKLAMPNTAN--NGPYTLKATWTKTGNPGRVAVLVPNSLW 689
Qy 715 SILDIRSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDLALGQGYRIISGGYSLGANSYF 774
Db 690 SILDIRSAHSAIQASVDGRSYCRGLWVSGVSNPFYHDDRDLALGQGYRIISGGYSLGANSYF 749
Qy 775 GSSMFGLAFTFVGRSKDYVYCRSNHACIGSVYLSQALCGSYLFGDAFIRASYGFGN 834
Db 750 GSSMFGLAFTFVGRSKDYVYCRSNHACIGSVYLSQALCGSYLFGDAFIRASYGFGN 809
Qy 835 QHMKTSYTPABESDVRWNNCLAGEIGAGLPVITPSKLYNELRPFVQABPYADHESPT 894
Db 810 QHMKTSYTPABESDVRWNNCLAGEIGAGLPVITPSKLYNELRPFVQABPYADHESPT 869
Qy 895 THEGDQARAPKSHLANLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETTL 954
Db 870 THEGDQARAPKSHLANLSVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGTETTL 929
Qy 955 HQETWTTTDAFLARHGVMVGRSGVSYASITSNIEYVGHGRYEYRDSRGYGLSAGSRVRF 1012
Db 930 HKETWTTTDAFLARHGVMVGRSGVSYASITSNIEYVGHGRYEYRDSRGYGLSAGSRVRF 987

RESULT 3
B86547
polymorphic outer membrane protein G family [imported] - Chlamydia pneumoniae (strain
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: B86547
R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise,
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: AB6491; MUID:20330349; PMID:10871362
A/Accession: B86547
A/Status: preliminary
A/Molecule type: DNA
```

A:Residues: 1-973 <STO>
A:Cross-references: UNIPROT:Q9Z896; UNIPARC:UPI000004707B; GB:BA000008; NID:98978824; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_13
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 973;
Best Local Similarity 33.5%; Pred. No. 5.8e-75;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY 1 MOTSFHKPFLSMILAYSCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPGTTVF 60
DB 1 MTSIRKPLISTLTAPCPAST---AFTVEVIMPSNFDDSGSKI-PPYTTLSDPRTGLCI 56
QY 61 SAGELTKNLDSIAALPLSCFNLGSLFTVLRGHSHTFENIRSTNGAALSNSAADG- 119
DB 57 FSGDLVIANLDAISRTSSCFNRRAGALQILKGGVFLNIRSGADGAISVITQNP 116
QY 120 ---LFTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNE 176
DB 117 ELCPISFSGFQWIFDNCESLT-----SDTSASNVIPIHSAIYATTPMLPTND 165
QY 177 KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTAOADGGACQVVTFSAMANEAPIA 236
DB 166 SILFOYNRSGAGFAAIGRTSITIENTKSLIFNGNGISNGGALTGSAAILNINSAPVI 225
QY 237 FVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGCVFNIRSGAGIYANGN 296
DB 226 FSTNATGIYGGAIYLT-----GGSMILTS-----GNLSGVLVFNNSRSGGAIYANGN 272
QY 297 VAPLNNKTLFLNNVASPV-YIAAKOPTSQOASNTSNYDGGAIKCKNGAQSNNSGS 355
DB 273 VTFNNSDLTTFQNTTASPNLSPPATPPPTTAVTP-LLGYGGAICTPPTPPPTGV-S 330
QY 356 VSPDGGVGFSSNVAAGKGAIYAKKLSVANGCPVQFLNINADGGAIVLGSSELSSLS 415
DB 331 LTISSGNSVTFLENIASEQGALYKKISIDNSKSTIFLNTAGKGAIAIPESGELSLS 390
QY 416 ADYGDIIIFDGNLKRKTAKENAAVNGVTVSSQAISSMGSGKITTILRAKAGHQLIFNDPIEM 475
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGATQGYTLFYDPI-- 439
QY 476 ANGNQPAQSSKLLKIN-----DGEYTGDIVP-----ANGSSTLYQNVITIE 517
DB 440 TSDDLGAASAAATVWVNPKASADG-AYSGTIVFSGETLTATEATPANATSTLNQKLELE 498
QY 518 QGRIVLRKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQPPQPPAANQILITLSNLSLS 576

A:Residues: 1-973 <STO>
A:Cross-references: UNIPROT:Q9Z896; UNIPARC:UPI000004707B; GB:AE001363; NI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_13
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 973;
Best Local Similarity 33.5%; Pred. No. 5.8e-75;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY 1 MOTSFHKPFLSMILAYSCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPGTTVF 60
DB 1 MTSIRKPLISTLTAPCPAST---AFTVEVIMPSNFDDSGSKI-PPYTTLSDPRTGLCI 56
QY 61 SAGELTKNLDSIAALPLSCFNLGSLFTVLRGHSHTFENIRSTNGAALSNSAADG- 119
DB 57 FSGDLVIANLDAISRTSSCFNRRAGALQILKGGVFLNIRSGADGAISVITQNP 116
QY 120 ---LFTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNE 176
DB 117 ELCPISFSGFQWIFDNCESLT-----SDTSASNVIPIHSAIYATTPMLPTND 165
QY 177 KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTAOADGGACQVVTFSAMANEAPIA 236
DB 166 SILFOYNRSGAGFAAIGRTSITIENTKSLIFNGNGISNGGALTGSAAILNINSAPVI 225
QY 237 FVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGCVFNIRSGAGIYANGN 296
DB 226 FSTNATGIYGGAIYLT-----GGSMILTS-----GNLSGVLVFNNSRSGGAIYANGN 272
QY 297 VAPLNNKTLFLNNVASPV-YIAAKOPTSQOASNTSNYDGGAIKCKNGAQSNNSGS 355
DB 273 VTFNNSDLTTFQNTTASPNLSPPATPPPTTAVTP-LLGYGGAICTPPTPPPTGV-S 330
QY 356 VSPDGGVGFSSNVAAGKGAIYAKKLSVANGCPVQFLNINADGGAIVLGSSELSSLS 415
DB 331 LTISSGNSVTFLENIASEQGALYKKISIDNSKSTIFLNTAGKGAIAIPESGELSLS 390
QY 416 ADYGDIIIFDGNLKRKTAKENAAVNGVTVSSQAISSMGSGKITTILRAKAGHQLIFNDPIEM 475
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGATQGYTLFYDPI-- 439
QY 476 ANGNQPAQSSKLLKIN-----DGEYTGDIVP-----ANGSSTLYQNVITIE 517
DB 440 TSDDLGAASAAATVWVNPKASADG-AYSGTIVFSGETLTATEATPANATSTLNQKLELE 498
QY 518 QGRIVLRKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQPPQPPAANQILITLSNLSLS 576

DB 830 SSRILNLKQIIPFVKAEVAYATHGGIOENTPEGRIFGHGHLNLAIVAVGVGFRGNHNRP 889
QY 929 NKYSFMAAYICDAVTTISGTETILLSHOETWTWDAFLHARGVWVVRGSMYASLTNIIEVY 988
DB 890 DFYTIIVAYADVTNRNPDCTTLPINGATWTSIGNNLTRSTLLVQASHTSVNDVLEIF 949
QY 989 GHGREYERDASRGYGLSAGSRVF 1012
DB 950 GHGCGDIRRTSRQVTLDIGSKLRF 973

RESULT 4

F72076
polymorphic outer membrane protein g family - Chlamydomophila pneumoniae (strain CWL029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: F72076
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: F72076
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <ARN>
A:Cross-references: UNIPROT:Q9Z896; UNIPARC:UPI000004707B; GB:AE001363; NI
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp_13
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 973;
Best Local Similarity 33.5%; Pred. No. 5.8e-75;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

QY 1 MOTSFHKPFLSMILAYSCSLNGGGYAAEIMVPGQIYDGETLTVSPFYTVIGDPGTTVF 60
DB 1 MTSIRKPLISTLTAPCPAST---AFTVEVIMPSNFDDSGSKI-PPYTTLSDPRTGLCI 56
QY 61 SAGELTKNLDSIAALPLSCFNLGSLFTVLRGHSHTFENIRSTNGAALSNSAADG- 119
DB 57 FSGDLVIANLDAISRTSSCFNRRAGALQILKGGVFLNIRSGADGAISVITQNP 116
QY 120 ---LFTIEGKELSPNCNLSLAVLPAATNKGSTPTTSTPSNGTIYSKTDLLLNNE 176
DB 117 ELCPISFSGFQWIFDNCESLT-----SDTSASNVIPIHSAIYATTPMLPTND 165
QY 177 KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTAOADGGACQVVTFSAMANEAPIA 236
DB 166 SILFOYNRSGAGFAAIGRTSITIENTKSLIFNGNGISNGGALTGSAAILNINSAPVI 225
QY 237 FVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEFDGCVFNIRSGAGIYANGN 296
DB 226 FSTNATGIYGGAIYLT-----GGSMILTS-----GNLSGVLVFNNSRSGGAIYANGN 272
QY 297 VAPLNNKTLFLNNVASPV-YIAAKOPTSQOASNTSNYDGGAIKCKNGAQSNNSGS 355
DB 273 VTFNNSDLTTFQNTTASPNLSPPATPPPTTAVTP-LLGYGGAICTPPTPPPTGV-S 330
QY 356 VSPDGGVGFSSNVAAGKGAIYAKKLSVANGCPVQFLNINADGGAIVLGSSELSSLS 415
DB 331 LTISSGNSVTFLENIASEQGALYKKISIDNSKSTIFLNTAGKGAIAIPESGELSLS 390
QY 416 ADYGDIIIFDGNLKRKTAKENAAVNGVTVSSQAISSMGSGKITTILRAKAGHQLIFNDPIEM 475
DB 391 ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFAITLGATQGYTLFYDPI-- 439
QY 476 ANGNQPAQSSKLLKIN-----DGEYTGDIVP-----ANGSSTLYQNVITIE 517
DB 440 TSDDLGAASAAATVWVNPKASADG-AYSGTIVFSGETLTATEATPANATSTLNQKLELE 498
QY 518 QGRIVLRKAKLSVNSLSQSGS-LYMEAGSTWDFVTPQPPQPPAANQILITLSNLSLS 576

Db 499 GGTALRNAGATLVNHFQDEKSVVIMDAGTT--LATNGANNITDGA---ITLNLKLVINL 553
Qy 577 SLLANNVTPPTNPPAQDSHPAIVG-STTAGSVTISGPFFEDLDDTAYDRYDMLGSN 635
Db 554 DSLDGTKA-----AVVNVQSTNGALTISGTLGLVKNQSDCCDNHGMFNKD 598
Qy 636 -QKINVLKQLGTGKPPANAPSDL-TLGNEMPKYGYGGSWKLAWDPTANNPPTLKATWT 693
Db 599 LQOVPILELKATSNVTITDPSLGTNGYQOSPYGQGTWETIDTT-----HVTGNWK 653
Qy 694 KTYNPGPERVASLVPNSLWGSILDIRSAHSAIOASVDGRSY-CRGLMVSGVSNFFYHDR 752
Db 654 KTYGLPHERLAPLIPNSLMANVIDLRAVSA--SAADGEDVPGKQLSITGITNFFHANH 711
Qy 753 DALQGGYRIYIGGYSLGANSY---FGSMFGLAFTEVFGRSKYVYVCRSNHHACTGSYVL 809
Db 712 TGDARSYRHMGGYLI--NTYTRITPDAAISLPGQPLTKSKDYLVLGHGSHNVYATVYS 769
Qy 810 STQOALCG-SYLFQDGAIRASVYFCGNHMKTSYTPABESDVVRDNNCLAGEIGLPIVI 868
Db 770 NITKSLFGSSRRFPFGGTRVYSRNEKVTSYTKLPKGRCSWSNCCWLGEGLNLPITL 829
Qy 869 TPSKLYLNELRPFVQAFPSYADHESFTEEGDQARAFKSGHLINLSPVGVKFDRCSSHP 928
Db 830 SRRILNLKQLIPPVKAEVAYATHGGIQENTPEGRIFGHLNLAIVAVPVGVAFGKSHNRP 889
Qy 929 NKYSFMAAYICDAYRTISGTTTLLSHOETWTDDAPHLARHGTVVVRGSMYASLTSNIRVY 988
Db 890 DPTIIVAVADPVYRHNPDCTTLTPINGATWTSIGNNLTRSTLLVQASSTSVNDVLEIF 949
Qy 989 GHGREYRDASRGYGLSAGSRVP 1012
Db 950 GHGCCDIRRTSRQYTLDIGSKLRF 973

RESULT 5
C81593
polymorphic membrane protein G family CP0299 [imported] - Chlamydomophila pneumoniae (stra
C/Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
R/Read: T.D.; Brubham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, M.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: C81593
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-995 <REA>
A/Cross-references: UNIPARC:UPI00001655F9; GB:AE002191; GB:AE002161; NID:g7189216; PIDN:
A/Experimental source: strain AR39, HL cells
C/Genetics:
C/Gene: CP0299
C/Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 25.9%; Score 1365.5; DB 2; Length 995;
Best Local Similarity 33.5%; Pred. No. 6e-75;
Matches 350; Conservative 171; Mismatches 420; Indels 103; Gaps 27;

Qy 1 MCTSHKFFLSMILAYCSLNGGYYAAIIVPQGIYDGEILTVSPFTVIGDPSGTTVP 60
Db 23 MKTSIRKFLISTLAPCPAST---AFTVEVIMPSNFDPSSGKI--FPYTTLSDPRGTLCI 78
Qy 61 SAGELTLKLNLSIALPLSCFNLGSLTFLVGRGHSITFENIRTSNTGAALSNSAAG- 119
Db 79 FSGDLYIANLDAISRTSSCSFNSRAGALQILKRGVGFNLIRSDAGALISSVITQNP 138
Qy 120 ---LFTIBGFKELSFNSCNLSLAVLPAATAATNKGSTPTTTTPSPNGTTYSKTDLLLANE 176
Db 139 ELCPLSPSGFGSOMIDNCESLT-----SDTSASNVIPHASAIYATTPMLFTNND 187
Qy 177 KFSFYSNLVGDGGDAIDAKSLTVQGISKLCVFOENTAOAGCQCVTSPSAMANEAFIA 236

Db 188 SILFOYNSAGFGAAIRGTSITIENTKKSLFNGNGSISNGGALTGSAAINLNNASVI 247
Qy 237 FVANVAGVRGGGIAAODGQGVSSSTSTEDPVVSPFRNTAVBFDGNVARVGGGYSYCN 296
Db 248 FSTNATGIYGAAYLT-----GGSLTIS-----GNLSGVLFVNNSSRGGGAIYANGN 294
Qy 297 VAFNLNGKTLFLNNVAGPV-YIAAKQPTSGQASNTSNNGYGGGAIFCQKGAQAGSNNGS 355
Db 295 VTFNNSDLTFQNTATSPQNSLPAPTPTPPVATP-LLGYGGAIFCTPTTPTPTPTGV-S 352
Qy 356 VSPQEGVVPFSSNVAACKGAIYAKLVANCPVQVPLRNIANDGGAIYLGESGELS 415
Db 353 LTISSGNSVTPLENIASEGGALYGGKISIDSNKSTIPLGNTAGKGAIAIPESGELS 412
Qy 416 ADYGDIIIPDGNLKTAKENAAVNGVTVSSQAIISWGGKITTIRAKAGHOILFNDPIEM 475
Db 413 ANQGDILFNKLSITSG-----TFTRNSIHFGKDAKPAITLGCATGYTLYFYDPI-- 461
Qy 476 ANGNQPAQSSKLLKIN-----DGEYTGDIIVP-----ANGSSTLYQNVITR 517
Db 462 TSDDLSSAASAAATVVPKASADG-AVSGTIVFSGETLTATEAATPANATSTLNQKLE 520
Qy 518 QGRIVLREKAKLVNSLSQSGS-LYMEAGSTWDPVTPQPPQPPAANQLITLSHLISL 576
Db 521 GGTALRNAGATLVNHFQDEKSVVIMDAGTT--LATNGANNITDGA---ITLNLKLVINL 575
Qy 577 SLLANNVTPPTNPPAQDSHPAIVG-STTAGSVTISGPFFEDLDDTAYDRYDMLGSN 635
Db 576 DSLDGTKA-----AVVNVQSTNGALTISGTLGLVKNQSDCCDNHGMFNKD 620
Qy 636 -QKINVLKQLGTGKPPANAPSDL-TLGNEMPKYGYGGSWKLAWDPTANNPPTLKATWT 693
Db 621 LQOVPILELKATSNVTITDPSLGTNGYQOSPYGQGTWETIDTT-----HVTGNWK 675
Qy 694 KTYNPGPERVASLVPNSLWGSILDIRSAHSAIOASVDGRSY-CRGLMVSGVSNFFYHDR 752
Db 676 KTYGLPHERLAPLIPNSLMANVIDLRAVSA--SAADGEDVPGKQLSITGITNFFHANH 733
Qy 753 DALQGGYRIYIGGYSLGANSY---FGSMFGLAFTEVFGRSKYVYVCRSNHHACTGSYVL 809
Db 734 TGDARSYRHMGGYLI--NTYTRITPDAAISLPGQPLTKSKDYLVLGHGSHNVYATVYS 791
Qy 810 STQOALCG-SYLFQDGAIRASVYFCGNHMKTSYTPABESDVVRDNNCLAGEIGLPIVI 868
Db 792 NITKSLFGSSRRFPFGGTRVYSRNEKVTSYTKLPKGRCSWSNCCWLGEGLNLPITL 851
Qy 869 TPSKLYLNELRPFVQAFPSYADHESFTEEGDQARAFKSGHLINLSPVGVKFDRCSSHP 928
Db 852 SSRILNLKQLIPPVKAEVAYATHGGIQENTPEGRIFGHLNLAIVAVPVGVAFGKSHNRP 911
Qy 929 NKYSFMAAYICDAYRTISGTTTLLSHOETWTDDAPHLARHGTVVVRGSMYASLTSNIRVY 988
Db 912 DPTIIVAVADPVYRHNPDCTTLTPINGATWTSIGNNLTRSTLLVQASSTSVNDVLEIF 971
Qy 989 GHGREYRDASRGYGLSAGSRVP 1012
Db 972 GHGCCDIRRTSRQYTLDIGSKLRF 995

RESULT 6
G86546
polymorphic outer membrane protein G family [imported] - Chlamydomophila pneumoniae (stra
C/Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86546
R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; MUID:20330349; PMID:10871362
A/Accession: G86546
A/Status: preliminary
A/Molecule type: DNA

A;Residues: 1-928 <STO>
A;Cross-references: UNIPROT:O9RB65; UNIPARC:UPI000002FFFO; GB:BA000008; NID:g8978821; PI
A;Experimental source: strain J138
C;Genetics:
C;Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 20.9%; Score 1101; DB 2; Length 928;
Best Local Similarity 30.5%; Pred. No. 6.4e-59;
Matches 324; Conservative 150; Mismatches 402; Indels 188; Gaps 34;

QY 1 MTSHPKFLSMILA-YSCCSLNGGGYAAEIMVPGIYDGETLVSPYTVIGDPSGTVV 59
DB 1 MKSQFSLVLSSTLACFTSCSTVFAA-TAENIGPSDSFDGSTNTGT--YTPKNTTIGIDY 57

QY 60 FSAGELTKNLDSNAALPLSCFNLGSPFVLGRHSLTPENIRTSNGALSNSAAG 119
DB 58 TLTGDTITLQNLGDS-AALTKGCFSDTTESLSPAGKGYSLFLNKSABGAALS-VTTDK 115

QY 120 LPTIEGKELSPSNCNLSLAVLPAATTNKGSTPTTTSPSNGTIYSKTDLLLNKFS 179
DB 116 NLSLTFGSSLT--LAAPSVITTPSG-----KGAKCGDLTFDNNGTIL 159

QY 180 FYSNLVSGDGGAIKSLTVQGISKLCVQENTAQDGGACQVWTSFSAANEAPFAVA 239
DB 160 FKQDYCEENGGAISTKNLSKNSTGSIPEGNKSSA-----SGDADVT 270

QY 240 NVAGVRGGIAAVDQGGQVSSSTSTEDPVVFSRNTA-VEFDGNVAR-VGGGIYSYGNV 297
DB 196 --TGKKGGAICA-----TGT-----VDITNTAPTFLSNIAEAGGAINSTGNC 238

QY 298 AFLNNGKTLPLNVAASPVVIAAKQTPSGQASNTSNNGYDGGAIKCKNGAQAAGNSGSYS 357
DB 239 TITGNTSLVFSN-----SVTAGNGGAL-----SGDADVT 270

QY 358 FDGEGVFPSSNVAAGKGAIAKGLSVAN--CGPVQFLRNI-----ANDGGAIIYLGESG 410
DB 271 ISGNQSVTFSGNQAVANGAIIYAKKLTASGGGGISFSNNIVQGTAGNGGAISILAAG 330

QY 411 ELSLADYGDIIIPDGN-----LKRTAKENADVNGVTVSSQAISMGSQKITTILRAKAGH 465
DB 331 ECSLSAEAGDITFNGNAIVATTPTTKRNSIDI-----GSTAKITNLRAISGH 378

QY 466 QILFNDPIEMANGNPOAOSKLLKINDGEG-----YTGDIVPA-----NGS 507
DB 379 SIFFYDPITA-----NTAADSTDTLNLNKADAGNSTDYSIVFSGEKLSDEKAVADNLT 434

QY 508 STLYQNVTIEQGRIVLRERAKLSVNSLSQT--GGSLYMEAGSTWDFVTPQPQPPAANQL 566
DB 435 STLKQPVTLTLAGNLVLRGVTLTKGFTQTAGSSVIMDAGTTL-----KASTE 483

QY 567 ITLSNLHLSSLILLANNAVTPNPPAQDHPAVIGSTTAGSVTISGPIFFEDLDITAY 626
DB 484 VTLTGLSIPVDSL-----GEGKKVVIASAASAKNVALSGPILLIDNQNAY 529

QY 627 DRYDMLGSQKINVLKLO-LGTXPPANASDLTLGNEMPKYQGSQKLAWDENTANNCP 685
DB 530 ENHD-LGKTQDFSFVQLSALGTATTTDVPAPVTATP--THYGYQGTWGMWDDTAST-P 586

QY 686 YTLKAT--WTKCYNPGPERVASLVNLSLWGSILDIRSAHSAIQASVDGRSVCRLWVSG 743
DB 587 KTKTATLWNTGYLNPENRQGPLVFNLSLWGSFSDIQAIGVTERALTLCSDRGFWAAG 646

QY 744 VSNFFYHRRDALQGGYRISGGYSL--GANSYFGSSMFLGATFVPRGSKDYVVCRSNHHA 802
DB 647 VANFLDKCKGKRYKHKSGGYAIGGAAQTCSENILISPAFCQLFGSDKDFLAKNHTDT 706

QY 803 CIGSVYLS-----TQOALCGSILFEDAFI--RASYGFGNQHKMTSYFAESDV 849
DB 707 YAGAFYIQTHTGCSFGTICLLDLKPLGSSWSHKPLVLEGLAYSHSVNSDLTKTYATPEVK 766

QY 850 RWDNNCLAGEIGAGLPVITPSKLY-LNELRPPVQAEFSYADHESFTTEEGDQARAFKSGH 908

DB 767 SWGNNAFNMMLGASHSY--PEYLHCFDPTVAPYIKLNLTYIRQDSFSEKTEGSRFDDSN 824

QY 909 LLNLSPVGVKDFRCSSSTHPNKYSFMAAYICDAYRTTISGTETTLSSHQETWTTDAFHAR 968
DB 825 LFNLSLPIGVKFEKFCDCNPSYDLTLISYVPLRNDPKCTTALVISGASMETVANNLAR 884

QY 969 HGVVVRGSMYASLTSTNIEVYGHGRYERDASRGYGLSAGSRVPF 1012
DB 885 QALQVRAGSHYAFSPMEFVLQGFVEVRGSSRIYNVDLGGKPFQF 928

RESULT 7
GB1591
polymorphic membrane protein G family CP0303 [imported] - Chlamydia pneumoniae (stra
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: G81591
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: AB1500; MUID:20150255; PMID:10684935
A;Accession: G81591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-928 <REA>
A;Cross-references: UNIPROT:O9RB65; UNIPARC:UPI000002FFFO; GB:AE002192; GB:AE002161; NII
A;Experimental source: strain AR39, HL cells
C;Genetics:
C;Gene: CP0303
C;Superfamily: Chlamydia pneumoniae polymorphic outer membrane protein G

Query Match 20.9%; Score 1101; DB 2; Length 928;
Best Local Similarity 30.5%; Pred. No. 6.4e-59;
Matches 324; Conservative 150; Mismatches 402; Indels 188; Gaps 34;

QY 1 MTSHPKFLSMILA-YSCCSLNGGGYAAEIMVPGIYDGETLVSPYTVIGDPSGTVV 59
DB 1 MKSQFSLVLSSTLACFTSCSTVFAA-TAENIGPSDSFDGSTNTGT--YTPKNTTIGIDY 57

QY 60 FSAGELTKNLDSNAALPLSCFNLGSPFVLGRHSLTPENIRTSNGALSNSAAG 119
DB 58 TLTGDTITLQNLGDS-AALTKGCFSDTTESLSPAGKGYSLFLNKSABGAALS-VTTDK 115

QY 120 LPTIEGKELSPSNCNLSLAVLPAATTNKGSTPTTTSPSNGTIYSKTDLLLNKFS 179
DB 116 NLSLTFGSSLT--LAAPSVITTPSG-----KGAKCGDLTFDNNGTIL 159

QY 180 FYSNLVSGDGGAIKSLTVQGISKLCVQENTAQDGGACQVWTSFSAANEAPFAVA 239
DB 160 FKQDYCEENGGAISTKNLSKNSTGSIPEGNKSSA-----SGDADVT 270

QY 240 NVAGVRGGIAAVDQGGQVSSSTSTEDPVVFSRNTA-VEFDGNVAR-VGGGIYSYGNV 297
DB 196 --TGKKGGAICA-----TGT-----VDITNTAPTFLSNIAEAGGAINSTGNC 238

QY 298 AFLNNGKTLPLNVAASPVVIAAKQTPSGQASNTSNNGYDGGAIKCKNGAQAAGNSGSYS 357
DB 239 TITGNTSLVFSN-----SVTAGNGGAL-----SGDADVT 270

QY 358 FDGEGVFPSSNVAAGKGAIAKGLSVAN--CGPVQFLRNI-----ANDGGAIIYLGESG 410
DB 271 ISGNQSVTFSGNQAVANGAIIYAKKLTASGGGGISFSNNIVQGTAGNGGAISILAAG 330

QY 411 ELSLADYGDIIIPDGN-----LKRTAKENADVNGVTVSSQAISMGSQKITTILRAKAGH 465
DB 331 ECSLSAEAGDITFNGNAIVATTPTTKRNSIDI-----GSTAKITNLRAISGH 378

QY 466 QILFNDPIEMANGNPOAOSKLLKINDGEG-----YTGDIVPA-----NGS 507
DB 379 SIFFYDPITA-----NTAADSTDTLNLNKADAGNSTDYSIVFSGEKLSDEKAVADNLT 434

QY 508 STLYQNVTIEQGRIVLRERAKLSVNSLSQT--GGSLYMEAGSTWDFVTPQPQPPAANQL 566


```
Db 435 STLKQPVLTAGNLVLRGVLDTKGTQTAGSSVIMDAGTTL-----KASTE 483
Qy 567 ITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAY 626
Db 484 VTLTGLSPVDSL-----GEGKKVIAAASAKNVALSGPILLDDNQWY 529
Qy 627 DRYDWLGSQKINVLKLO-LGTKPPANAPSDLTLGNEMPKYQYQSWKLANWDPNTANNP 685
Db 530 ENHD-LGKTQDFSFVQLSALGATATTDDPAVPTVATP--THYGYQGTWGTWVDDTAST-P 586
Qy 686 YTLKAT--WTKTYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCGLWVG 743
Db 587 KTKATLAWTGYLPNERQGLVPNSLWGSFSDIQAIOGVIERSAULTCSDRGFMAAG 646
Qy 744 VSNFFYHRRDALGQGYRISGYSI-LGANSYFGSSMFLAFTEVFGRSKDYVVCRSNHA 802
Db 647 VANFLDKKKBKRYKHSKGYAIGGAACQCSENLISPAFCQLFGSKDFLVAKNHTDT 706
Qy 803 CIGSVLS-----TQALCGSYLFGDAFI--RASYGFGNQHMKTSYTFABESDV 849
Db 707 YAGAFYIQHITECSGFIOCLLDKLPFSWSHKPLVLEGLAYSHVSNLDLTKYTAPEVK 766
Qy 850 RWDNNCLAGEIGAGLPVITPSKLY-LNELRPVQAFPSYADHESFTEEGDQARAFKSGH 908
Db 767 SWGNAPFNMQLGASHSY--PEYLHCFDYAPYIKNLITYIRQDSFSEKGTGRSFDSDN 824
Qy 909 LLNLSPVGVKFDRCSTTHPKNSFMAAYICDAYRTISGTETTLTSHOETWTTDAFHAR 968
Db 825 LFNLSLPIGVKFEKSCDNDPSYDLTSLVVPDLIRNDPKCTALVISGASMETYANNLAR 884
Qy 969 HGTVRGWYASLTNSINIEYHGRIYEDRDSRGYGLSAGSRVP 1012
Db 885 QALQVRAGSHYAFSPMFVQLQFVPEVRGSRIRYVNDLGGKQF 928

RESULT 8
D72077
polymorphic outer membrane protein g family - Chlamydomphila pneumoniae (strain CWL029)
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: D72077
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:9920606; PMID:10192388
A:Accession: D72077
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <ARN>
A:Cross-references: UNIPROT:086164; UNIPARC:UPI000002PFEP; GB:AE001628; GB:AE001363; NID
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp 11
C:Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G

Query Match 20.7%; Score 1090; DB 2; Length 928;
Best Local Similarity 28.8%; Pred. No. 3e-58;
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;

Qy 1 MQTSFHKPFLSNILAYSC--CSLGGGYAAINWVPGIYDGETLTVSF-----PYTVI 51
Db 1 MKTSPFWLVSVLAFAFSLQSLAN-----ELLSPDSDFNQINDSGTFTPKTSATYSLT 56
Qy 52 GDPSTTVFVSGAGELTKNLNLSIAALPL--SCFNGLLGSFTVLGRGHSLTPTENRTSTNG 109
Db 57 GD--VPPYEPKGT-----PLSDSCFKQTTNLFNLGNHSLTFGPDAGTHA 102
Qy 110 AALNSAADGLFTIEGPKELSFNSCNLAVLPAATYNGSQTPPTTSPNGTIYSKTD 169
Db 103 GAAASTANKNLTFSGFSLSPDS-----SPETTIVTQGTLLSAGG 144
Qy 170 LLLLNNEKFSYSNLVSGDGGDAIDAKSLTVQGISKLVQFQNTAQDAGCAQCVTSPSAM 229
```

```
Db 145 VNLNIRKLVVAGNFSFADGGAIKGASFLTGTSGDALFSNNSSSTKGA----- 194
Qy 230 ANEAPIAFVANVAGVRGGIAAQQGGVSSSTSTEDPVVFSRNTA--VEFDGNVARVG 288
Db 195 -----IATTAGAR-----IANNITYVFLSNIASTS 220
Qy 289 GGIYSYGNVAPFLNNGKTLPLNNVASPVVIAAKQPTSGOASNTSNYGGGAIKCKNGAQA 348
Db 221 GGAIDDEGTSILSNKELYP-----EGNAAKTT-----GGAI-CNTRKASG 259
Qy 349 G-----SNNGSVSPDGBGVFPSSNVAAGKGGAIYAKLSVANCGPQVQFLRNTAND--- 400
Db 260 SPELIISNNK-----TLIFASNVAFTSGGAHAKKALSGALNPGYQGTILFSGETLTADBLK 310
Qy 401 -GGAIYIGESGELSADYGDIIIPDGNLKRKTAKENAADVNGVTVSSQAISSMGSGKITTLL 459
Db 311 KGGAISIDASGELSLSAETGNITFVRNLTITT--GSTD---TPKRNAINIGSNKFTTEL 364
Qy 460 RAKAGHQLPNDPIEMANGNNQPAQSKLLKINDGEG-----YTGDIVFA----- 504
Db 365 RAAKNHTIFFYDPI-----TSEGTSSDLKINNGSAGALNPYQGTILFSGETLTADBLK 418
Qy 505 ---NGSSTLYQNVTIEGRIVLRKAKLSVNSLSQTCGSLY--WEAGSTWDPVTPQPPQP 560
Db 419 VADNLKSFQPVLSGGKLLQGVLTLESTFSFQSEAGSLGMDSGITLSTTAGS----- 473
Qy 561 PAANQLITLSNLHLSLSLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFED 620
Db 474 -----ITITNLGINVDSLGLKQPV-----SLTAKGASKNKIVVSGKLNLD 513
Qy 621 LDDTAYDRYDWLGSQKINVLKLOGLTKPPANAPSDLTTLGNEMP-----KTYGQGSW 672
Db 514 IEGNIYBSH-MFSDHQLFSLIKITV---DADVTNVDISSLIPVPAEDPNSEYGFQGO 568
Qy 673 KLANDPNTANNGPYTLKATWTKGYNCPGPERVASLVPNSLWGSILDIRSAHSAIQASVDG 732
Db 569 NVNWTDTATN-TKEATATWTKGTFVSPSPERKSALVCNTLWGVFTDIRSLOQLVEIGATG 627
Qy 733 RSYCRGLWVSGVSNFFYHRRDALGQGYRISGYSIAGANSYF-GSSMFLAFTVFGRSK 791
Db 628 MEHKQGVFVSSMTNPLHKTGDNKRGFRHTSGGIVIGGSAHTPKDDLFTFAFCHLPARDX 687
Qy 792 DYVVCNRNHHACIGSVYLSLTOQAL-----CGSYLFGDAFI-----RASYG 831
Db 688 DCPTAHNRSRTYGTLPFKHSHTLQPNYLRGRKAFSESAIEKFPREIPLALDVQVSFS 747
Qy 832 PGNOHMKTSYTFAESDVRWDDNCLAGEIGAGLPVITPSKLYLNLRLPFPVQAFSPYADH 891
Db 748 HSDNRMETHYTSLPESGSGWSNECIAAGIGLDLPVLNSNPHLPKFTPIPMKVMVYVSG 807
Qy 892 ESFTEEGDQARAFKSGHLLMLSVPGVKFDRCSSTHNPKYSFMAAYICDAYRTISGTETT 951
Db 808 NSPFESSDDGFGSIGRLNLSIPVGAKFVQGDIGDSYTYDLSGFFVSDVYRNNPQSTAT 867
Qy 952 LLSHOETWTTDAFHARHGVVVVRGVSASVTSNTIEVGHGRIYEDRDSRGYGLSAGSRV 1011
Db 868 LVMSPDGKIRGNGNLRSQAFLRGSNNYVNSNCELFGHYAMELRGSSRNTYNDVGTCLR 927
Qy 1012 P 1012
Db 928 P 928

RESULT 9
H86546
polymorphic outer membrane protein G family [imported] - Chlamydomphila pneumoniae (strain H86546)
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
```

A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: H8546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <STO>
A:Cross-references: UNIPROT:O86164; UNIPARC:UPI000002FFBF; GB:BA000008; NID:g8978822; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp 11
C:Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Query Match 20.7%; Score 1090; DB 2; Length 928;
Best Local Similarity 28.8%; Pred. No. 3e-58;
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;

QY 1 MCTSFHKPFLSMILAYSC--CSLNGGYYAAEIMVPGIYDGETLTVSF-----PYTVI 51
DB 1 MKTSPWLVSSVLAFLSCLQSLAN-----EELSPDDSFNGNIDSGTFTPKTSATYSLT 56

QY 52 GDPSTGVTSAGELTLKLNLDNSIAALPL--SCFNLGSGFTVLGRGHSITFENIRTSNG 109
DB 57 GD---VFFYEPGKGT-----PLSDSCPKQTDDNLTLPLNGHSLTFFGIDAGTHA 102

QY 110 AALNSAAGDLFTIEGFKELSPNCNLSLLAVLPAATNKGSOPTTTPSPNGTIYSKTD 169
DB 103 GAAASTTANKNLTFSGFSLSPDS-----SPSTVTVTGQTLSSAGG 144

QY 170 LLLLNNEKFSFYENLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTSFSAM 229
DB 145 VNLENIRKLVVAGNPFSTADGGAIKGASFLLTGTSGDALFNSNSSTKGA----- 194

QY 230 ANEAPIAFVANVAGVGGGIAAVQDGOQGVSSSTSTEDPVVFSRNTA-VEFDGNVARVG 288
DB 195 -----IATTAGAR-----IANNVGYRFLSNIASTS 220

QY 289 GGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKOPTSQASNTNNGYDGGAIKCKNGAQA 348
DB 221 GGAIDDEGTSILSNKFLYF-----EGNAAKTT-----GGAI-CNTKASG 259

QY 349 G-----SNNSGVSFDGEGVFFSSNVAGKGAIAVAKLSVANCQVPQFLRNIAND--- 400
DB 260 SPELIIISNNK-----TLIFASNVAETSGGAIHAKKLALSGGGTPEFLRNVSSATP 310

QY 401 -GGAIVLGESELSLSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAISMGGSKITTL 459
DB 311 KGGAIISIDASGELSLSAETGNITFVRNLTATT--GSTD-----TPKRNAINIGSKPTFL 364

QY 460 RAKAGHOILPNDPLEMANGNQPAAQSSKLLKINDGEG-----YTGDIVPA----- 504
DB 365 RAAKHHTIFFYDPT-----TSBGTSDVILKINNGSAGALNPYQGTILFSGEITLTADELK 418

QY 505 ---NGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTGGSLY-MEAGSTWDPVTPPQPQP 560
DB 419 VADMLKSPFQPSVLSGKLLQKGVLTSTSFQEAGSLGMDSGTTLSTTAGS----- 473

QY 561 PAANQLITLNLHLNLSLANNVATNPPTNPPAQDHPAVIGSTTAGSVTISGPIPFED 620
DB 474 -----ITITNLGINVDLSGLKQPV-----SLTAKGASKNIVTSGKMLNID 513

QY 621 LDDTAYDRYDLWLSGNQKINVLKLOCTKPPANAPSDTLGNEMP-----KYVGQGSW 672
DB 514 IEGNIYESH-MFSDHQLFSLKITYV-----DADVTNVDISLIPVPAEDPNSEYGFQGW 569

QY 673 KLAWDPTNANGPVTLKATVTKGYNPQPERVASLVPNSLWGSILDIRSAHSATQAQSDVG 732
DB 569 NVNMTTDTATN-TKEATATVTKTGFEVSPERKSALVCNLTNGVTDIRSLQQLVEIGATG 627

QY 733 RSYCRGLWVSGVSNFFYHHRDALQGGVRYITSGGYSGLANSYF-GSSMFGLAFTVPGORSK 791
DB 628 MEHQGFVWSMTNLFHKTGDNKRGFRHTSGGVVIGGSAHTPKDDLFTFAFCHLFARDK 687

QY 792 DYVVCRSNHHACIGSVYLSLTOQAL-----CGSYLFGDAFI-----RASYG 831

DB 688 DCFIAHNNSRTYGGTLPFKHSHTLOPQNYLRLGRAPSESAIEKPPREIPLALDVQVSFS 747
QY 832 PCNQHMKTSYTFABESDVRWNNCLAGIGAGLPVITPFSKLYLNLNLPFPVQAEFSYADH 891
DB 748 HSDNEMETHYISLPESEGSWNECIAGGIGLDLPFLVSLNPHPLPKTFIPQMKVEMVYSQ 807
QY 892 ESFTBEGDOARAPKSGHLLNLISVPVGVKPDRCSSSTHPNKYSFMAAYICDAYTIGSTETT 951
DB 808 NSFPSSSDGRGFSIGRLINLSIPVGAKFVQGDIGDSYTDLSGPFVSDVYRNNPQSTAT 867
QY 952 LLSHQETWTTDAFLARHGVVVRGSMYASLTSENIYVGHGRVEYDASRGYGLSAGSRVR 1011
DB 868 LVMSPDWMKIRGNLSRQAFLLRGSNNVYNSNCLFGRHAMELGRSSRYNVVDVGTCLR 927
QY 1012 F 1012
DB 928 F 928

RESULT 10
F81591
polymorphic membrane protein G family CP0302 [imported] - Chlamydothila pneumoniae (str
C:Species: Chlamydothila pneumoniae; Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
C:Accession: F81591
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: F81591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-949 <REA>
A:Cross-references: UNIPARC:UPI00001655FA; GB:AE002192; GB:AE002161; NID:g7189226; PIDN
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0302
C:Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Query Match 20.7%; Score 1090; DB 2; Length 949;
Best Local Similarity 28.8%; Pred. No. 3.1e-58;
Matches 311; Conservative 154; Mismatches 394; Indels 222; Gaps 31;

QY 1 MCTSFHKPFLSMILAYSC--CSLNGGYYAAEIMVPGIYDGETLTVSF-----PYTVI 51
DB 22 MKTSPWLVSSVLAFLSCLQSLAN-----EELSPDDSFNGNIDSGTFTPKTSATYSLT 77

QY 52 GDPSTGVTSAGELTLKLNLDNSIAALPL--SCFNLGSGFTVLGRGHSITFENIRTSNG 109
DB 78 GD---VFFYEPGKGT-----PLSDSCPKQTDDNLTLPLNGHSLTFFGIDAGTHA 123

QY 110 AALNSAAGDLFTIEGFKELSPNCNLSLLAVLPAATNKGSOPTTTPSPNGTIYSKTD 169
DB 124 GAAASTTANKNLTFSGFSLSPDS-----SPSTVTVTGQTLSSAGG 165

QY 170 LLLLNNEKFSFYENLVSGDGAIDAKSLTVQGISKLCVFOENTAQADGGACQVTSFSAM 229
DB 166 VNLENIRKLVVAGNPFSTADGGAIKGASFLLTGTSGDALFNSNSSTKGA----- 215

QY 230 ANEAPIAFVANVAGVGGGIAAVQDGOQGVSSSTSTEDPVVFSRNTA-VEFDGNVARVG 288
DB 216 -----IATTAGAR-----IANNVGYRFLSNIASTS 241

QY 289 GGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKOPTSQASNTNNGYDGGAIKCKNGAQA 348
DB 242 GGAIDDEGTSILSNKFLYF-----EGNAAKTT-----GGAI-CNTKASG 280

QY 349 G-----SNNSGVSFDGEGVFFSSNVAGKGAIAVAKLSVANCQVPQFLRNIAND--- 400
DB 281 SPELIIISNNK-----TLIFASNVAETSGGAIHAKKLALSGGGTPEFLRNVSSATP 331

QY 401 -GGAIVLGESELSLSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAISMGGSKITTL 459

A81591
polymorphic membrane protein G family CP0307 [imported] - Chlamydomophila pneumoniae (strain C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81591
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.P.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-930 <REA>
A:Cross-references: UNIPROT:Q92393; UNIPARC:UPI00001655PB; GB:AE002193; GB:AE002161; NID
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0307
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 20.6%; Score 1083; DB 2; Length 930;
Best Local Similarity 28.4%; Pred. No. 8e-58;
Matches 304; Conservative 160; Mismatches 406; Indels 202; Gaps 29;

QY 1 MQTSFHKPFLSMILAYSCSLNGGGYAAEI-MVPGIYDGETLTVSPFVTVIGDPSGTVV 59
DB 1 MKIPLHKLISSTLV-TPILLSIATYGADASLSPDSDPGAGSGTFTPKST-ADANGTNY 58

QY 60 PSAGELTLKNDNSIAALPLSCFGLNLLGSFTVLGRHSLTFENIRTSNGAALSAAADG 119
DB 59 VLSGNNYINDAGKG-TALTGCCFTTGTDLTFTGKGYSFNTVDAGSNAGAAASTADK 117

QY 120 LFTIEGFKELSPNCSNLAVLPAATNKGSTPTTT---STPSNGTIYSKTDLLLNNE 176
DB 118 ALTTFTGFSNLSP-----IAAPGTTVAGSKTSLSAGALNTDNGTILFSONV----- 164

QY 177 KFSFYNLVSGDGAIDAKSLTVQGISKLCVQFNTAQADGGACQVTVTSFSAWANEAPTA 236
DB 165 -----SNEANNNGGAIITKTLISGNTSITFTSNKAKLGGAIYSSAAASISGNTGQLV 219

QY 237 FVANVAGVGGGIAAVQDQGGQSSSTSTEDPVVFSRNTAVEFDGNVARVGGIYSYGN 296
DB 220 FMNN-KGETGGGALGFE-----ASSITQNSLSPFGNTATD----- 255

QY 297 VAFNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGGGAIKCKNGAQAGSNNGSV 356
DB 256 -----SHTIN-----QAFTQPLVVPVTAATAASDIYIDALLTSFVQTPPEHYGQHW 582

QY 676 W-DPNTANNGPYTLKATWTKTGPNPGRVVASLVPNSLNGSILDIRSAHSAIQASVDGRS 734

DB 583 WADTSTAKSG-----TMTVTTGPNPERRASVVPDLSWASFTDRTLQQTMTSOANSIY 638
QY 735 YCRGLWVSGVSNPFYHDDRDALGCGQYRISGYSGLANSY-PGSMFGLAPTEVFGSKDY 793
DB 639 QORGLWASGTANFPHKDKSGTNQAFRHKSIGYIYVGGSAEDSENFISVAFQOLFGKDKDL 698
QY 794 VVCSNHHACIGSVYLSSTQQAQCSYL--FGDAF-----IRASYFGNQHMKTS 840
DB 699 FIVENTSINYLASYLQHRAFLGGLPMPSPGSIITDMLKDIPLILNAQLSYSTTKNDMDTR 758
QY 841 YTPAESDVRWDDNCLAGEIGAGLPIVITPCKLYLNELRPVQAEFSVADHESFTEEGDQ 900
DB 759 YTSYPEAQSGMTNNSGALGGLSLALYLPKEAPFFQGYFPFLKQAVYSROONPKESGAE 818
QY 901 ARAPKSGHLLNLSVPVGVKPKDRCSSTHNKYSFMAAICDAVRTISGTETLLSHQETWT 960
DB 819 ARAFDGDLVNCSPVGIKLEKISEDEKNNFEISLAYIGDVYRKNPRETSLSLWVGASWT 878
QY 961 TDAFHARHGVVVRGSMYASLTNIYVGHGRYEDASRGYGLSAGSRVRF 1012
DB 879 SLCKNLARQAFLASAGSHLTLSPHVELSGEAAVELRGSAAHYNVDCGLRYSF 930

RESULT 13
D72078
polymorphic outer membrane protein g family - Chlamydomophila pneumoniae (strain CML029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: D72078
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72078
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-930 <ARN>
A:Cross-references: UNIPROT:Q92393; UNIPARC:UPI000004708C; GB:AE001627; GB:AE001363; NID
A:Experimental source: strain CML029
C:Genetics:
A:Gene: pmp_8
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 20.5%; Score 1079; DB 2; Length 930;
Best Local Similarity 28.3%; Pred. No. 1.4e-57;
Matches 303; Conservative 160; Mismatches 407; Indels 202; Gaps 29;

QY 1 MQTSFHKPFLSMILAYSCSLNGGGYAAEI-MVPGIYDGETLTVSPFVTVIGDPSGTVV 59
DB 1 MKIPLHKLISSTLV-TPILLSIATYGADASLSPDSDPGAGSGTFTPKST-ADANGTNY 58

QY 60 PSAGELTLKNDNSIAALPLSCFGLNLLGSFTVLGRHSLTFENIRTSNGAALSAAADG 119
DB 59 VLSGNNYINDAGKG-TALTGCCFTTGTDLTFTGKGYSFNTVDAGSNAGAAASTADK 117

QY 120 LFTIEGFKELSPNCSNLAVLPAATNKGSTPTTT---STPSNGTIYSKTDLLLNNE 176
DB 118 ALTTFTGFSNLSP-----IAAPGTTVAGSKTSLSAGALNTDNGTILFSONV----- 164

QY 177 KFSFYNLVSGDGAIDAKSLTVQGISKLCVQFNTAQADGGACQVTVTSFSAWANEAPTA 236
DB 165 -----SNEANNNGGAIITKTLISGNTSITFTSNKAKLGGAIYSSAAASISGNTGQLV 219

QY 237 FVANVAGVGGGIAAVQDQGGQSSSTSTEDPVVFSRNTAVEFDGNVARVGGIYSYGN 296
DB 220 FMNN-KGETGGGALGFE-----ASSITQNSLSPFGNTATD----- 255

QY 297 VAFNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNYGGGAIKCKNGAQAGSNNGSV 356
DB 256 -----SHTIN-----QAFTQPLVVPVTAATAASDIYIDALLTSFVQTPPEHYGQHW 582

QY 676 W-DPNTANNGPYTLKATWTKTGPNPGRVVASLVPNSLNGSILDIRSAHSAIQASVDGRS 734

```
Db 275 TISGNKSLTPAENS SVTGGALCAUGLDLSAAGPTLFNNRCNTAAAGKGAIAIADSGS 334
Qy 412 LLSLSADYGDIIIPDGNLKR TAKENAAADVGVTVSSQAISMGGSGKITTLIRAKAGHQILFND 471
Db 335 LLSLSANQGDITPLGN-----TLTSTAPTSTRNAIYLGSSAKITNLRAAQGQSIIYVD 387
Qy 472 PIEMANGN-----NQPAQSGKLKINDGEGYTGDIYPA-----NGSSTLY 511
Db 388 FIASNTTGASDVLITINQPDNSPL-----DYSCTIVFSGEKLSDAKAADNFSTILK 440
Qy 512 QNVTEQGRIVLREKAKLSVNSLSOTGGSLYMEAGSTWDFVTPQPPQPPAANQLITLSN 571
Db 441 QPLALASGTALKGNVLDVNGFTTEGSTL-----MQGTGKLKADTEAISLTK 490
Qy 572 LHLSSLSLANNNAVTPNPPAODSHPAVIGSTTAG---SVTISGRPIFFEDLDDTAYDR 628
Db 491 LVVDLSALBGNKSV8-----IETAGANKTITLSPVFDSSGNPYE- 532
Qy 629 YDLGSGNOKINVLKQLGTGP-----PANAPSDLTLG-----NEMPKYGYQGSWKL 675
Db 533 -----SHTIN---QAFTQPLVWFPTAATAASDIYIDALLTSPVQTPPEHYGYQGHWEAT 582
Qy 676 W-DPNTANNPVTLKAATKTGYNPGRVRSVSLVNSLWGSTLDIRSAHSAIQLASVDCRS 734
Db 583 WADTSTKSG---TWMTVTGYNPNERRASVDPDSLWASFTDRTIQQIMTSQANSIY 638
Qy 735 YCRGLWVGVSNIFFYHRRDALQGGYRIYISGGYSLGANSY-FGSSMFGLAFTVFGRSKDY 793
Db 639 QORGLWASGTANFFHKDKSGTNQAPRHKS YGVIYVGS AEDFSENIFSVAFQCLFGKDKDL 698
Qy 794 VVCRNHHACTGSVYLSLQQALCGSYL--FGDAP-----IRASYGFGNQHMKTS 840
Db 699 FIVENTSHYLAASYLQHRALFGLGMPSPSGSITDMLKDIPILNLAOLSYSYTKNDMDTR 758
Qy 841 YTPAESDVRWNNCLAGEIGAGLPVITPCKLYLNELRPVQAFSADHESFTEEDQ 900
Db 759 YTSYPEAGSMTNNGSALGELGSLALYLPKGAFFQGGYFPFLFOAVYSRQONPKESGAE 818
Qy 901 ARAFKSGHLLMLSVPGVKFORCSSTHPNKYSFMAAYICDAYRTTSGTETILLSHQRTWT 960
Db 819 ARAFDGDLVNCISIPIVGIRLEKISDEKNNFEISLAYIGDVYRKNPRSTRISLWVGASWT 878
Qy 961 TDAFHARHGVVVRGSMVASITSNTEVYGHGRYEVDRASRGVGLSAGSRVRF 1012
Db 879 SLCKNLARQAFLASAGSHLTLSPHVELSGEAAVELRGSAAHYNVDGILRYSF 930

RESULT 14
C72078
polymorphic outer membrane protein g family - Chlamydothila pneumoniae (strain CML029)
C1species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C1Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C1Accession: C72078
R1Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A1Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A1Reference number: A72000; MUID:99206606; PMID:10192388
A1Accession: C72078
A1Status: preliminary
A1Molecule type: DNA
A1Residues: 1-936 <ARN>
A1Cross-references: UNIPROT:Q92988; UNIPARC:UPI00000470BC; GB:AE001627; GB:AE001363; NID
A1Experimental source: strain CML029
C1Genetics:
A1Gene: pmp 7
C1Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Query Match 19.8%; Score 1041; DB 2; Length 936;
Best Local Similarity 29.0%; Pred. No. 2.9e-55;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

Qy 1 MQTSFHKFFLSNLAISCCSLNGGGYAAIMV--PQGIYDGETLTVSPFYTVIGDPSGTT 58
```

```
Db 1 MKSSVSLWFSSIPFSSLSI-----VAAEVTLDSNNNSYDGSNGTTFTVTFSTTDAAGTT 56
Qy 59 VFSAGELTLKNLDSIAALPLSCFNLGSLFTVLGRGHSITFFENIRT-STNGAALSNSAA 117
Db 57 YSLSDSVSPQNALGALPIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGTVASTSA 116
Qy 118 DGLPTIIRGFKELSFNSCNLSLAVLPAATNKGSTPTTTSTPNSNGTYSKTDLLLNNEK 177
Db 117 DKNLLFDFSLSIISCPSLIL-----SPT-----GQCALKSVGNLSLTGNSQ 159
Qy 178 PSFYSNLVSGDGGDAIDAKSLTVQGISKLCVQENTAQADGGACQVTVSFSMAWEAPAP 237
Db 160 IIFTQNFSSDGGVINTKNFLSLGTSQFASPRNQA----- 195
Qy 238 VANTVAGVRGGIAAIVODQQQVSSSTSTEDP-VVSFSRNTAVZFDGNVARVGGGIYSYN 296
Db 196 ---FTGQGGVYVA-----TGTITIENSPOIIVSPQNLA---KGS---GGALYSTDN 238
Qy 297 VAFNNGKTLPLANNVSPVYIAAKQPTSGQASNTSNNYGOGGAI FCKNGAAGANNSSGV 356
Db 239 CSITDNFQVIFDGNSA---WEAAQ-----AOGGAICC-----TTTDKTV 274
Qy 357 SPDCGCVVPFSSNVAAGKGAIAKLSVANCVPQVFLNLTAND-----CGALYLGSSG 410
Db 275 TLTGNKLSFTNTALTGGGALSGLKVSISAGGPTLFQSNISGSSAGGGGGAIAIASAG 334
Qy 411 ELSLSADYGDIIIPDGNLKR TAKENAAADVGVTVSSQAISMGGSGKITTLIRAKAGHQILFN 470
Db 335 ELALSATSGDITFNN-----QVTNGSTSTFNAINIIDTAKVTSIRATQSIYFY 385
Qy 471 DPEMANGNQPAGSSKLLKINDGEG-----YTGDIVF-----ANGSSTLYQ 512
Db 386 DPIT---NPGTAASDTLNLNLADANSEIYGAIVFSGEKLSPTEKATAAANTVTSTIRQ 441
Qy 513 NVTIEQGRIVLREKAKLSVNSLSOTGGS-LYMEAGSTWDFVTPQPPQPPAANQLITLSN 571
Db 442 PAVLARGDLVLRDGVTVTFKDLTQSGSRILMDGGTTLS-----AKEANLSNG 490
Qy 572 LHLSSLSLANNNAVTPNPPAODSHPAVIGSTTAGSVTISGRPIFFEDLDDTAYDRYDW 631
Db 491 LAVNLSSLDGTNKA-----ALKTEAADKNISLGTIALIDTEGSEYENHN- 535
Qy 632 LGSNQKINVLKQLGTGTPPAN-----APSDLTIGNEMPKYGYQGSWKLAWDPNTANNGP 685
Db 536 LKSASTYPLLEL---TTAGANGTITLGAALSTLTQEBETHYGYQGNQMSWANATSSK-- 590
Qy 686 YTLKATWTKGYNPGRVRSVSLVNSLWGSTLDIRSAHSAIQLASVDCRSYCRGLWVGVS 745
Db 591 -IGSINWTRGTGYPSPRKNLPLNSLWGNPIDIRSNQLIETKSGEPFERELWLSGIA 649
Qy 746 NFFYHRRDALQGGYRIYISGGYSLGANSYF-GSSMFGLAFTVFGRSKDYVVCRRNHHACI 804
Db 650 NFFYRDSNPTFHGRFHSISGGYALGITATPAEDQLTFAPCOLFARDNRNHI TGNHGDYTG 709
Qy 805 GSVYLSLQQAL--CGSYLFGDAP-----IRASYGFGNQHMKTSYTFE 844
Db 710 ASLYFHTHEGLFDIANFLWKGATRAPWVLSISQIIPLSFPAKPSYLTIDNHMKTYTT-- 767
Qy 845 EESDVR--WNNCLAGIBIGAGLPVITPCKLYLNELRPVQAFSADHESFTEEDQAR 902
Db 768 DNSIISGSRNDACADLGASLPVIVSYL-LKEVEFPVKVQYIAHQDQFYRHAEGR 826
Qy 903 AFKSGHLLNSVPGVKFDRCSSTHPNKYSFMAAYICDAYRTTSGTETILLSHQRTWTD 962
Db 827 AFNKSSELINVEIPIGVTFPERDSKSEKGYDILTWYILDAYERNPKCKTSLIASDANMAY 886
Qy 963 AFHLARHGVVVRGSMVASITSNIEVYGHGRYEVDRASRGVGLSAGSRVRF 1012
Db 887 GTNLARQGSVRAAHNFQVPHMFIQGFAPFVSSSRNTNLTGSKFCF 936
```

RESULT 15

B81591
polymorphic membrane protein G family CP0308 [imported] - Chlamydomophila pneumoniae (stra
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81591
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: B81500; MUID:20150255; PMID:10684935
A:Accession: B81591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-936 <REA>
A:Cross-references: UNIPROT:Q92898; UNIPARC:UPI0000131CED; GB:AE002193; GB:AE002161; NID
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0308
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 19.7%; Score 1040; DB 2; Length 936;
Best Local Similarity 29.0%; Pred. No. 3.3e-55;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MOTSPHKFPLSMILAYSCCLNGGGYAAEIMV--PQGIYDGETLTVSPFVTVIGDPSGTT 58
DB 1 MKSSVSWLPFSSIPLFSSLSI-----VAAEVLDSNNSSYDGSNGTTFVFTTDAAGTT 56

QY 59 VFSAGELTKNLNDNSTAALPLSCFNLGSLTFVLGRHSLTPENIRT-STNGGAALNSAA 117
DB 57 YSLSDVSFQNALGIPPLASGCFLEAGDGLTFQGNQHALKFAFINAGSAGTVASTSAA 116

QY 118 DGLFTTEGFKLSFNSCNLSLLAVPLAATNKGSGTPTTTPSNGTIYSKTDLLLNNEK 177
DB 117 DKNLLFNDPFRSLIISCPSLLL-----SPT-----GQCALKSVGNLSLTGNSQ 159

QY 178 PSFVSNLVSDGGAIDAKSLTVQISKLVPQENTAAQDGAQGVVTSFSAANEAPIAF 237
DB 160 IITQNFSSDNGGVINTKNFLSGTSGQSFASFRNQA----- 195

QY 238 VANVAVRGGGIAAQQDQGGVSSSTSTEDP-VVSFSRNTAVPDGNVAVRGGGIYSYGN 296
DB 196 ---FTGQGGVVYA-----TGTITIENSPGIVSFQNLA---KGS---GGALYSTDN 238

QY 471 DPIEMANGNNQPAOSSKLLKINDGEG-----YTGDIVF-----ANGSSTLYQ 512
DB 386 DPIT---NPTAASTDTLNLADANSEIEYGAIVFSGEKLSPTEKATAANVTSTIRQ 441

QY 513 NVITTEQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPPPQPPAANQLITLSN 571
DB 442 PAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDGGTLLS-----AKEANLSING 490

QY 572 LHLSSLSILANNAVTPNTPPAQDSDHPAVIGSTTAGSVTISGPIPFEDLDDTAYDYDW 631
DB 491 LAVNLSSLDGNTKA-----ALTEAADKNISLSGTHIALIDTSGSYFENH- 535

QY 632 LGSNOKINVLQLQGTGKPPAN-----APSDLTGLNEMPKYGYQGSWKLAWDPTNANGP 685
DB 536 LKSASTYPLLEL---TTAGANGTITLGLSTLTLEQPEPTHYGYQGNQLSWANATSSK-- 590

QY 686 YTLKATWTKTGNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLGVSGVS 745

DB 591 -IGSNWTRTGYIPSPERKSNLPLNSLWGNFIDIRSNQLIETKSSGPPFERELWSGIA 649
QY 746 NFFYHDDRALCOQVRYISGYSGLGANSYF-GSSMFGLAFTVFGRSKDYVWVCRNHHACI 804
DB 650 NFFYRDSMPTRHGFHISGGFALGATATTPAEDQLTFAPCQLFADRNRHITCKNHDYVG 709
QY 805 GSVYLSLTOQAL--CGSYLFGDAF-----IRASYGFGNQMKTSYTFPA 844
DB 710 ASLYPHHTEGLFDIANFLWGLKATAPWVLSISQIIPLSFDAKPSYLTNDNMKTYTT-- 767

QY 845 BESDVR--WDNNCLAGEIGAGLPIVITPSKLYLNLRLRPVQAEPSYADHESPTBEGDQAR 902
DB 768 DNSIIKGSWRNDAFAFCADLGASLFPVISVPYL-LKEVEFPVKVQYIYAHQDDFYERYAEGR 826

QY 903 APKSGHLNLNLVPVGVKDRCSSTHPNKYSFMAAYICDAYRTISGETTLLSHQETWTTD 962
DB 827 AFNKSELINVEIPIGVTTERDSKSEKGYDILMTYLDAYRENPKCOTSLIASDANMWAY 886

QY 963 AFHLARHGVVVRGSMYASLTSNIEVYHGRVEYRDASRGYGLSAGSRVRF 1012
DB 887 GTNLARQGSVRAAHNFQVNPHEIFGQFAPEVRSSRNNTNLGSKFCF 936

RESULT 16
C86546
polymorphic outer membrane protein G family [imported] - Chlamydomophila pneumoniae (stra
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I-
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: C86546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-936 <STO>
A:Cross-references: UNIPROT:Q92898; UNIPARC:UPI0000131CED; GB:BA000008; NID:98978817; P
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp 7
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 19.7%; Score 1040; DB 2; Length 936;
Best Local Similarity 29.0%; Pred. No. 3.3e-55;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MOTSPHKFPLSMILAYSCCLNGGGYAAEIMV--PQGIYDGETLTVSPFVTVIGDPSGTT 58
DB 1 MKSSVSWLPFSSIPLFSSLSI-----VAAEVLDSNNSSYDGSNGTTFVFTTDAAGTT 56

QY 59 VFSAGELTKNLNDNSTAALPLSCFNLGSLTFVLGRHSLTPENIRT-STNGGAALNSAA 117
DB 57 YSLSDVSFQNALGIPPLASGCFLEAGDGLTFQGNQHALKFAFINAGSAGTVASTSAA 116

QY 118 DGLFTTEGFKLSFNSCNLSLLAVPLAATNKGSGTPTTTPSNGTIYSKTDLLLNNEK 177
DB 117 DKNLLFNDPFRSLIISCPSLLL-----SPT-----GQCALKSVGNLSLTGNSQ 159

QY 178 PSFVSNLVSDGGAIDAKSLTVQISKLVPQENTAAQDGAQGVVTSFSAANEAPIAF 237
DB 160 IITQNFSSDNGGVINTKNFLSGTSGQSFASFRNQA----- 195

QY 238 VANVAVRGGGIAAQQDQGGVSSSTSTEDP-VVSFSRNTAVPDGNVAVRGGGIYSYGN 296
DB 196 ---FTGQGGVVYA-----TGTITIENSPGIVSFQNLA---KGS---GGALYSTDN 238

QY 471 DPIEMANGNNQPAOSSKLLKINDGEG-----YTGDIVF-----ANGSSTLYQ 512
DB 386 DPIT---NPTAASTDTLNLADANSEIEYGAIVFSGEKLSPTEKATAANVTSTIRQ 441

QY 513 NVITTEQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPPPQPPAANQLITLSN 571
DB 442 PAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDGGTLLS-----AKEANLSING 490

QY 572 LHLSSLSILANNAVTPNTPPAQDSDHPAVIGSTTAGSVTISGPIPFEDLDDTAYDYDW 631
DB 491 LAVNLSSLDGNTKA-----ALTEAADKNISLSGTHIALIDTSGSYFENH- 535

QY 632 LGSNOKINVLQLQGTGKPPAN-----APSDLTGLNEMPKYGYQGSWKLAWDPTNANGP 685
DB 536 LKSASTYPLLEL---TTAGANGTITLGLSTLTLEQPEPTHYGYQGNQLSWANATSSK-- 590

QY 686 YTLKATWTKTGNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSVCRLGVSGVS 745


```
Db      827 NVSVPIGTFERFSPNRASAEATVIYVADVVRKPNDCCTALLINNTSWKTTGTLSRQA 886
      QY      971 VVVRGSMYASLTNIEVYGHGRIYEDASRGYLSAGSRVRF 1012
      Db      887 GIGRAGIFYAPSPNLEVTNLSMEIRGSSRSYNADLGKGFQF 928

RESULT 18
B86546
polymorphic outer membrane protein G/I family [imported] - Chlamydomophila pneumoniae (str.
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: B86546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <STO>
A:Cross-references: UNIPROT:Q9Z398; UNIPARC:UPI0000047087; GB:BA000008; NID:98978819; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_9
C:Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 19.4%; Score 1021; DB 2; Length 928;
Best Local Similarity 29.5%; Pred. No. 4.7e-54;
Matches 313; Conservative 161; Mismatches 404; Indels 184; Gaps 39;

QY      1 MQTSPHKPFLSMILVYSCSLNGGGVAAEIMV---PQGIYDGE-TLTVSPPTVTVGDPGG 56
      Db      1 MKSSLHWFLISSSLALP-LSLNFSAPAAVVEINLPTNFSFGPTVT---PQAQTTNADG 56
      QY      57 TTVFSGAGELTKNLNDSIAALPLSCFNLGSPFVLGRGSLTFENIRFTSTGAALNSGA 116
      Db      57 TIYNLTGDSITNA-GSPALTATSCFKETGNLSFGQHGQFLLQWIDAGAN-CTPTNTA 114
      QY      117 ADGLFTIEGPKELSPNCNSLLAVLPAATNKGSTPTTSTPSNGTIYKTDLLLNNE 176
      Db      115 ANKLSFGFSYLSL-----LQTNATYTGAKSTGACISQSNY 154
      QY      177 KFSFYNLVSGDGAIDAKSLTVQGISKLCVFOENTQAQDGACQVVTFSAMANEAPIA 236
      Db      155 SCYFGQNFNDNGALQGSSISL-SLNPLNLPKAKATQKGGAL-----YST----- 200
      QY      237 FVANVAGVGGGTAQVQDGGQGVSSSTSTEDPVVPSRNTAVZFDGNVAVRGGIYSYN 296
      Db      201 -----GGIT-----INNTLS-----ASFSENTA-----ANNNGAIYTEAS 231
      QY      297 VAPLNGKTL-FLNNVASPVYIAAKQPTSCQASNTSNYDGGAIFFCNGAQAAGSNNGS 355
      Db      232 -SPISNKALSFNN-----SVTATSAI-----GGAIYC-----STSPAPKV 268
      QY      356 VSPDGGVGVFPFSSNAAGKGAIYAKKLSVANCGPVQFLNIAND-----GGAIYGEHG 410
      Db      269 LTLSDNGELNFIQNTAITSGAIYTDNLVSSGGPTLFKNNSAIDTAAPLGGAIATADSG 328
      QY      411 ELSLSADYGDIIIPDGNLKPTAKENADVNGVTVSSQAISNG-SGGKITTLRAKAGHQLIF 469
      Db      329 SLSLSALGGDITTEGN---TVVKGAS--SSQTTRNSINIGTNNAKIVQLRASQGNITYF 383
      QY      470 NDPIEMANGKNQPAQSSKLKINDGE-----GYTGDIVPA-----NGSSTLY 511
      Db      384 YDPII-----TSITAAOLDALNNGPDLAGNPVQGTIVFSGEKLSEAAEAEDNLSITIQ 439
      QY      512 QNVTFIEQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPPQPQPPAANQLITLS 570
      Db      440 QPLTLAGGQLSLKSGVTLVAKFSQSPGSLTLLMDAGITLETADG-----ITIN 487
      QY      571 NLHLSLSLLANNAVTPNPPAQDSDHPAVIGSTTAG-SVTISGPIFFEDLDLTAYDRI 629
```

```
Db      488 NLVLNVDLSL-----KETKATLKATQASQTVTLSSGSLSLVDPFSGNVYEDV 532
      QY      630 DWLGSNQKINVLQLQGTGKPPANAP-SDLT---LGNEMPKYGYGSGKWLANDPNTANNGP 685
      Db      533 SW--NNPQV-FSCULTLADDDPANIHITDLAADPLEKQPIHMGYQGNWALSQEDTATKSK 589
      QY      686 YTLKATWTKTCYNPGRVASLVNSLWMSGSLDIRSAHSAIAQASVDGRSVCRLWVSGVS 745
      Db      590 -AATLTWTKYGNPNPERGLTVANTLWGSFVDVRSIQQLVATKVRSQETRGWCEGIS 648
      QY      746 NFFYHRRDALQGVYRISGGYSLGANSYFGS-SMPGLAFTVEFGRSKDYVVCRRSHHACI 804
      Db      649 NFFHKDSTKINKFRHISAGYVVGATTTLASDNLITAAFCQLFGKDRDHFINKRASAYA 708
      QY      805 GSVLSTQOALCG-----SYLFG-----DAFIRASVGFQGNQHMKTSYTFABESDVRW 851
      Db      709 ASLHLQHLATLSSPSLLRYLFGSSSEQPVLPDAQI--SYIYSKNTMKTYITQAPKGSWSW 766
      QY      852 DNNCLAGBIGAGLPVITPFSKLYLNLNLPFPVQAFSADHESFTBEGDQ-ARAPKSGHLL 910
      Db      767 YDGCNLELASSLPHALTSHGLPHAYFPFIKVEASYIHQDSFKERNNTTLVRSFDSGLI 826
      QY      911 NLSVPGVGVKFCRCSSTHPNKYSFMAAYICDAYRTISGTETILLSHQETWTTDAPHLRHG 970
      Db      827 NVSVPIGTFERFSPNRASAEATVIYVADVVRKPNDCCTALLINNTSWKTTGTLSRQA 886
      QY      971 VVVRGSMYASLTNIEVYGHGRIYEDASRGYLSAGSRVRF 1012
      Db      887 GIGRAGIFYAPSPNLEVTNLSMEIRGSSRSYNADLGKGFQF 928

RESULT 19
B86546
polymorphic outer membrane protein G/I family [imported] - Chlamydomophila pneumoniae (str.
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86546
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: B86546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1276 <STO>
A:Cross-references: UNIPROT:Q9Z899; UNIPARC:UPI0000131CEC; GB:BA000008; NID:98978816; P:
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_6

Query Match 19.2%; Score 1012; DB 2; Length 1276;
Best Local Similarity 25.9%; Pred. No. 2.6e-53;
Matches 339; Conservative 153; Mismatches 417; Indels 398; Gaps 40;

QY      38 DGETLTVSPPTVIGDPSCTTVFSAGELTKNLNDSIAALPLSCFNLGSPFVLGRGHS 97
      Db      36 NGSSGGAAPFAKETS DASGTTTTLSDVSITNV-SAITPADKSCFTNTGALSFGVADHS 94
      QY      98 LTFENIRTSNNGAALNSAADGLFTIEGPKELSPNCNSLL----- 138
      Db      95 LVLQTLTALTHDGAANNVT-----ALSFSGFSSLLIDSAFATGTSGKGAIQCVT 144
      QY      139 -----AVLPA 143
      Db      145 NTEGGTATFDNASVTLQKNTSEKGAASVAYSIDLAKTTTAALLDONTSTKNGGALCST 204
      QY      144 ATI-----NGSQOT-PTTSTPSNGTIYS----- 166
      Db      205 ANTVQNSGTVTFSSNTATDKGGGYSKEKSTLDANTGVVTFKSNKTAKTGAWSSDDN 264
      QY      167 -----KTDLILLNNEKFSFYSNL 184
```

Db 265 LALTGNTQVLFOENKTTGSAQAANNPEGCGGAIACCYLATATDKTGLAISQNMSPFTSNT 324
Qy 185 VSGDGGAIIDAKSLTVQGISKLCVFOENTAQAD-GGACQVVT-SFSAMANEAPAFVANVA 242
Db 325 TTANGGAIYATKCTLDGNTTL-TFDQNTATAGCGAIYTETEDFSLKGSTGTVTFSTNTA 383
Qy 243 GVRGGI-----AAVDGQGV-----SSSTST----- 265
Db 384 KT-GGALYSKGNLSLTGNTNLLFSGNKATGPSNSANQEGCGGAILSFLESASVSTKGL 442
Qy 266 ---EDPVVFSRNTA-----VEFDGNVAR-VGGGIY----- 292
Db 443 WIEDENVSLSGNTATVSGGAIYATKCALHGNNTLTFDGNATAGGAIYTETEDFTLTG 502
Qy 293 SYGNVAFNN-----GKTLFLNNVASPVYIAAKOPTSQASNTSNYGDGGAIF 341
Db 503 STGTVTFTNTAKTAGALHTKNTSFTKNKA---LVPSGNSATATATTTTQEGCGGAIL 559
Qy 342 C-----KNGAQAGNSNGSVS-----FDGBGVVFPSSNVAACKGG 376
Db 560 CNISESDIATKSLTLENESLSFINNTAKRGGGIYAPKCVISGESINFPGNTAETSGG 619
Qy 377 AIYAKKLSVANCGPVQFRLNIANDGGAIYLGESGELSADYGDIIIPDGNLKRKTAKENAA 436
Db 620 AIYSKNLSITANGPVSFNNSGGKGAIIYADSGELSLEADGDITPSGN---RATEGTS 676
Qy 437 DVNGVTVSSQAIISMGGSGKITTILAKAGHOILFNDPIEM---ANG----- 478
Db 677 TPN-----SIHLGAGAKITKLAAPGHTIYFYDPTIMEAPASGGTTEELVINPVVKAI 729
Qy 479 -----NNQPAQSSKLIKINDGEGYTDIVF-----ANGSSTLYQNVTIEQG 519
Db 730 VPPQPKNGPIASVPVVPVAPANPTGTIVFSSGKLPQDASIPANTTTILNQKINLAGG 789
Qy 520 RIVLRKAKLSVNSLSQGG-S-LYNEAGSTWDFVTQPQQPPAANOLITLSNLHLSLS 578
Db 790 NVVLKEGATLVYSFTQPDSTVPMDAGTTLTETTN-----NTDGSIDLKLSVNLDA 843
Qy 579 LLANNAVTPNTPPAQDSHPAVIGTTPAGSVTISGPFPEDLDTDAYDRYDMLGSNQKI 638
Db 844 LDGKRMIT-----IAVNSTSGGLKISGDLKPHNNGSPYDN---POLKANL 886
Qy 639 NVLKLQL-GTKPPAN-----APSDLTIGNEMPKYQGSWKLANDPNTANNGPYTLKAT 691
Db 887 NLPFLDLSSSTSGTVNLDFFNPISMA-----APDYGQGSWTLV---PKVGAAGKVTLVAE 940
Qy 692 WTKTGYNPGRKVASLVNSLWGSTLDIRSAHSAIQASVDCGRSLVWSGVSNFFVHD 751
Db 941 WQALGYTPKPELRATLVNSLWYVNIHSTQOEIATAMSDAPSHPGIWIIGIGIAGAFQD 1000
Qy 752 RDALGQGVYISGGYSLGANSPGSS-MPLGAPTEVPGRSKDYVYCRSNHHACIGSVYLS 810
Db 1001 KOKENAGPRLISRGYIVGSGMTTPQEYTFVAFBSQLFGSKDYVVSIDKSQVYAGS--- 1056
Qy 811 TQQALC--GSYLP-----GDAPI-----RASYGNGQNMKTSYFAB 845
Db 1057 ----LCAQSSXVIFLHLSLRHLVSKVLPFLPGTPLVLHGQVSYGRHHNMITKLANNT 1112
Qy 846 ESDVRWDMNCLAGETAGLPVITPSKLYLNLAPFPVQAFPSYADHSFTEGEGDQAFK 905
Db 1113 QKSDWDHSHFAVEVGSLPVDL--NRYLTSYSPYVKLVVSVNQKQFQVAADPRIFD 1170
Qy 906 SGHLNLNVSPVGVKFDRCSSSTHPNKYSFMAAYICDARTISGTETTLSSHQETTTWTDAPF 965
Db 1171 ASHLVNVSIPLMGLFKHESAKPPSALLLLTGYADVADYDRHPCLTS-LTNGTSSMTFATN 1229
Qy 966 LARGVTVVGRSMYALSTNIEVYHGCRYEYDADRGCYGLSAGSRVF 1012
Db 1230 LSRQAFABASGHLKLHLGLDCFASGSCETLSRSSRSYNANCGTYSF 1276

RESULT 20
C81591

polymorphic membrane protein G family CP0309 [imported] - Chlamydophila pneumoniae (stra
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81591
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81591
A;Status: preliminary
A;Molecule type: DNA
A;Releides: 1-1276 <REA>
A;Cross-references: UNIPROT:Q92899; UNIPARC:UPI0000131CEC; GB:AE002193; GB:AE002161; NID
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0309

Query Match 19.2%, Score 1012, DB 2, Length 1276;
Best Local Similarity 25.9%, Pred. No. 2.6e-53;
Matches 339; Conservative 153; Mismatches 417; Indels 398; Gaps 40;
Qy 38 DGETLTVSPVTVIGDPSTGTTVFSAGBLTLKNLDNSTAALPLSCFGLGSLFTVLGRGHS 97
Db 36 NGSGSAAPFAKETSDASGTYTLTSDVSIINV-SAITPADKSCFTTGGALSFPVGADHS 94
Qy 98 LTFENIRTSNGAALSNAADGLFTTRGPKELSFNCNSLL----- 138
Db 95 LVLQTLIALTHDGAANNNTNT-----ALSPFGSFLILDSAPATGTSGGKAICVT 144
Qy 139 -----AVLPA 143
Db 145 NTEGGTATFDNASVTLQKNTSEKOGAAVSAYSIDLAKTTTAAALLDQNTSTKNGALCST 204
Qy 144 ATT-----NKGSTP-PTTSTPSNGTIYS----- 166
Db 205 ANTVQNGSVTVTFSSNTATDKGGIYSKEKSDLDANTGVTVFKSNATKTGGNWSDDN 264
Qy 167 -----KTDLLLLNNEKPSFYSNL 184
Db 265 LALTGNTQVLFOENKTTGSAQAANNPEGCGGAIACCYLATATDKTGLAISQNMSPFTSNT 324
Qy 185 VSGDGGAIIDAKSLTVQGISKLCVFOENTAQAD-GGACQVVT-SFSAMANEAPAFVANVA 242
Db 325 TTANGGAIYATKCTLDGNTTL-TFDQNTATAGCGGAIYTETEDFSLKGSTGTVTFSTNTA 383
Qy 243 GVRGGI-----AAVDGQGV-----SSSTST----- 265
Db 384 KT-GGALYSKGNLSLTGNTNLLFSGNKATGPSNSANQEGCGGAILSFLESASVSTKGL 442
Qy 266 ---EDPVVFSRNTA-----VEFDGNVAR-VGGGIY----- 292
Db 443 WIEDENVSLSGNTATVSGGAIYATKCALHGNNTLTFDGNATAGGAIYTETEDFTLTG 502
Qy 293 SYGNVAFNN-----GKTLFLNNVASPVYIAAKOPTSQASNTSNYGDGGAIF 341
Db 503 STGTVTFTNTAKTAGALHTKNTSFTKNKA---LVPSGNSATATATTTTQEGCGGAIL 559
Qy 342 C-----KNGAQAGNSNGSVS-----FDGBGVVFPSSNVAACKGG 376
Db 560 CNISESDIATKSLTLENESLSFINNTAKRGGGIYAPKCVISGESINFPGNTAETSGG 619
Qy 377 AIYAKKLSVANCGPVQFRLNIANDGGAIYLGESGELSADYGDIIIPDGNLKRKTAKENAA 436
Db 620 AIYSKNLSITANGPVSFNNSGGKGAIIYADSGELSLEADGDITPSGN---RATEGTS 676
Qy 437 DVNGVTVSSQAIISMGGSGKITTILAKAGHOILFNDPIEM---ANG----- 478
Db 677 TPN-----SIHLGAGAKITKLAAPGHTIYFYDPTIMEAPASGGTTEELVINPVVKAI 729
Qy 479 -----NNQPAQSSKLIKINDGEGYTDIVF-----ANGSSTLYQNVTIEQG 519
Db 730 VPPQPKNGPIASVPVVPVAPANPTGTIVFSSGKLPQDASIPANTTTILNQKINLAGG 789

QY 520 RIVLREKAKLSVNSLSQTGGS--LYMEAGSTWDFVTPOPPQQPPAANQLITLSNLHSLSS 578
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 790 NVVLKEGATLOVYSFTQQPDSTVFMDAGTLETTITTN-----NTDGSIDLKNLSVNLD 843
:
QY 579 LLANNAVTPNPAPPADSHPAVIGSTTAGSVISGPFIFFEDLDDTDAYRDYDWLGSNQKI 638
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 844 LDGKRMIT-----IAVNSTGGKLISGDLPKHNNEGSFYN---PGLKAML 886
:
QY 639 NVLKLOL-GTKPPAN-----APSDLTGLNEMPKYGQGSKWLAWDPNTANNPGYTLLKAT 691
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 887 NLPFLDLSTSGTVNLDDFNPIESSMA----APDYGYQSSTLV--PKVGAGGKVTLVAE 940
:
QY 692 WTKTYGNPGERVASLPNSLWSGIIDIRSAHSALQASVDGRSYCRGLWVSGVSNPFYHD 751
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 941 MQALGYTPRDELATLVPNSLNWAYNIHSIQEIATAMSDAPSPHGIMGTGGIGNAFHQD 1000
:
QY 752 RDALGOQRYVISGGYISLGANSYFGSS-MFGLATEVFPGRSKDYVVCESHHACIGSVYLS 810
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 1001 KOKENAGFRILSRGIYVGGSMITTFQEYTFAVAFSQLFGKSQDVVSIDIKSQVYAGS---- 1056
:
QY 811 TQQAIC--GSYLF-----GDAPT-----RASYPGNGOHMKTSYTFAE 845
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 1057 -----LCAQSYVIPLHSSLRRHVLSKVLPGLPETPLVLHGQVSYGRNHHTWKLANNT 1112
:
QY 846 ESDVRDNNCIAGEIAGLPIVITPSKLYLNEULRPVQAEFSYADHESFTTEGDQARAFK 905
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 1113 QGKSDWDSSHFAVEVGGSLPVDL--NYRYLTSYSPYVKLVQVSNOKGFQEVAAADPRIFD 1170
:
QY 906 SGHLNLSPVGVKPRDCSTHPNKYSFMAAYICDAVRTISCTETILLSHOEWTTTDAFH 965
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 1171 ASHLNVNYSIPMLGTLTFKHESAKPPSALLTLTGAYDAYDRHPHCCLTS-LTNGTSMSTPATN 1229
:
QY 966 LARHGVVVRSWASLTSTNEVGHGRYEYRDSRGYGLSAGSRVRF 1012
:|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 1230 LSRAQFPABASCHLKLHLGLDCFPASGCSELRSSRSYNACGRYSP 1276
:

RESULT 21
B72078
polymorphic outer membrane protein g/i family - Chlamydoiphila pneumoniae (strain C)
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: B72078
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimm Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:95206606; PMID:10192388
A:Accession: B72078
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1407 <ARN>
A:Cross-references: UNIPROT:Q9Z899; UNIPARC:UPI0000047086; GB:AE001627; GB:AE00 A:Experimental source: strain CWL029
C:Genetics:
A:Gene: pmp 6

Qy	209	QENTAQDGCACQVVTSPFSAANEAPIAVANVAGVRCGGIAAVQDGGQGVSSSTSTEDP	268
Db	584	SGNTATVSGGA--IYATKCALHGNTTLTDPGNTAETAGGAI-----YIETEDP	629
Qy	269	VVSFSRNTAEVDPGNVARVGGGIIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQAS	328
Db	630	TLTGSTGT-VTFSTNTAKTAGALHTKGNSTFTKNKALVFGNSA-----TATAT	677
Qy	329	NTSNNTGDCGALFC-----KNGAAGSNNSGSVS-----PDGEGV	363
Db	678	TTTTDQGGCGAILCNTISESDIATKSLTLENHSLSPINNTAKRSGGIIYAPKCVISGES	737
Qy	364	VPSSNVAAGKGAIYAKKLSVANCGPQVFLRNINDGGAIIYLGSEGLSLSADYGDIIIF	423
Db	738	INFQNTAETSGAIIYSKNLSITANGPVSFTNNSCGKGAIYIADSGELSLAIDGDITF	797
Qy	424	DGNKRTAKENAADVNGVTVSQAISMGSGGKIITLRKAKHQIILFNDPIEM---ANG--	478
Db	798	SGN---RATEGTSTPN-----SIHLGAKAKITKLAAPGHTIYFYDPIITMEAPASGGT	847
Qy	479	-----NNQPAQSSKLLKINDGEGYTGDIVP-----ANG	506
Db	848	IBELVINPVVKAIVPPQPKNGPIASVVPVPAPANPNTGTIIVFSSGKLPSQDASIPANT	907
Qy	507	SSTLYONVITBOGRIVLRREKAKLSVNSLSQTGS--LYMEAGSTWDFVTPOPPQPPAANQ	565
Db	908	TTTILNQKINLAGNVVLKEGATLOVYSFTQQPDSFTVPMDAGTTLETTTIN-----	961
Qy	566	LITLSNLHLSLSLILANNAVNTPTPPAQDSHPAVIGSTTAGSVIISGPIFFEDLDDTA	625
Db	962	SIDLKSLSVNLDAIDGKRMIT-----IAVNSTSGCLKIKSGDLKPHNNEGSF	1007
Qy	626	YDRYDWLGSNQKINVLKQLQ--GTKPPAN-----APSDLTLGNEMPKYGGQSGWKLAOWP	678
Db	1008	YDN---PELKANLNLPLDLSSTSGTVNLDDFPNIPSSMA-----APDYGYQSSWTLV--P	1058
Qy	679	NTANNGPYTLKATWTKTYGPNPGERVASLVPNSLWMSILDIRSAHSAIQASVDGRSVCRG	738
Db	1059	KVGAGGKVTLVAEQALGYTPKPELRATLVPNSLWNAVYNIHSITQOBIATAMSDAPSHPG	1118
Qy	739	LWISGVSNPFFHDBDALGQGVYITSGYSGYSGANSYFGSS--MFLAFTVEVGRSKDYVCCR	797
Db	1119	IWIGIGNAPHQDKOKENAGFRLISRGYIIVGGSWTTPQETTFVAVFSQLFKSKDYVUSD	1178
Qy	798	SNHHCATGVSVYLSQQALC--GSYLF-----GDAFI-----RASXYGF	832
Db	1179	IKSQVIYAGS-----LCAQSSYVIPLEHSLRHLVSKVLPELPGETPLVLHGQVSYGR	1230
Qy	833	GNQHKTSYTPFAESDVRWNONCLAGIAGLPLIVITPSKLYLNELRPPFQAERFSYADHE	892
Db	1231	NHNWNTKLANNTQCKSDWDSHPAVEVGGSLPVDL--NYRYLTSYFPYVKLQVVSYNQK	1288
Qy	893	SFTEGDOARAFKSGHLLNLVSPGVGVKDFDRCSSTHPNKYSFMAAYICDARTISGTETTL	952
Db	1289	GFOEVAADPRIFDASHILVNVSIPMGLTFTKHESAKPPSALLTLTGYADVADRHPCLITS-	1347
Qy	953	LSHQBETWTTDAFLARHGTVVVRGWSYASLTSNIEVYGHGRYERDASRGYGLSAGSVRPF	1012
Db	1348	IVTQSTCTSTATNTSPQAPAFASGHIKJHGLDPCPASGCELRSSRSRYNANGCTSYSF	1407

RESULT 22

H86492 Fmp_3 [imported] - Chlamydophila pneumoniae (strain J138)
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86492
E:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: H86492
A:Status: preliminary

A:Molecule type: DNA

A;Residues: 1-772 <STO>

A:Cross-references: UNIPROT:Q9RB71: UNIPARC:UPI000000CCC33: GB:BA000008: NID:q8978389: PT

A: Experimental source: strain J138

C: Genetics:

A:Gene: pmp 3 2

Query Match 18.6% Score 978.5 DB 2: Length 772:

Best Local Similarity 29.4%: Pred. No. 1.4e-51:

DONOR LOCAL SIMILARITY	25:74%	FREQ:	NO: 1:78-317
Matches	260:	Conservative	142:
Mismatches	260:	Indels	167:
		Gaps	26:

Qv 180 FYSNLVSGDGGAI DAKSLTVQGT SKLCVFOENTAOADGGACOVVTSFSAMANEAPTAFVA 239

3 PSKNPSTDNGGAI TAKTSLTGTMSALESENTSSKKGAIOTSDALTITGNOGEVSP-- 60

240 NVAGVRCGGTAAVODGOOGVSSSTSTEDPWSPSRNTAVPEQGNVARVCGGTYSVGNVAF 299

db 61 S P N T S S D S G A A T E T E A S V T T 80

300 L N N G K T T . E T . N N V A S P V Y T A A K O P T S G O A S N T S N N V G D G G A T P C K N G A O A G S N N G S V S E D 359

nb 81 SNNAKVSTPN-----KMGASSSTGDM-SGGATCAVK-----TSTNTKATTT 123

007 360 0801/0802/0803/0804/0805/0806/0807/0808/0809/0810/0811/0812/0813/0814/0815/0816/0817/0818/0819/0820/0821/0822/0823/0824/0825/0826/0827/0828/0829/0830/0831/0832/0833/0834/0835/0836/0837/0838/0839/0840/0841/0842/0843/0844/0845/0846/0847/0848/0849/0850/0851/0852/0853/0854/0855/0856/0857/0858/0859/0860/0861/0862/0863/0864/0865/0866/0867/0868/0869/0870/0871/0872/0873/0874/0875/0876/0877/0878/0879/0880/0881/0882/0883/0884/0885/0886/0887/0888/0889/0890/0891/0892/0893/0894/0895/0896/0897/0898/0899/0900/0901/0902/0903/0904/0905/0906/0907/0908/0909/0910/0911/0912/0913/0914/0915/0916/0917/0918/0919/0920/0921/0922/0923/0924/0925/0926/0927/0928/0929/0930/0931/0932/0933/0934/0935/0936/0937/0938/0939/0940/0941/0942/0943/0944/0945/0946/0947/0948/0949/0950/0951/0952/0953/0954/0955/0956/0957/0958/0959/0960/0961/0962/0963/0964/0965/0966/0967/0968/0969/0970/0971/0972/0973/0974/0975/0976/0977/0978/0979/0980/0981/0982/0983/0984/0985/0986/0987/0988/0989/0990/0991/0992/0993/0994/0995/0996/0997/0998/0999/1000/1001/1002/1003/1004/1005/1006/1007/1008/1009/1010/1011/1012/1013/1014/1015/1016/1017/1018/1019/1020/1021/1022/1023/1024/1025/1026/1027/1028/1029/1030/1031/1032/1033/1034/1035/1036/1037/1038/1039/1040/1041/1042/1043/1044/1045/1046/1047/1048/1049/1050/1051/1052/1053/1054/1055/1056/1057/1058/1059/1060/1061/1062/1063/1064/1065/1066/1067/1068/1069/1070/1071/1072/1073/1074/1075/1076/1077/1078/1079/1080/1081/1082/1083/1084/1085/1086/1087/1088/1089/1090/1091/1092/1093/1094/1095/1096/1097/1098/1099/1100/1101/1102/1103/1104/1105/1106/1107/1108/1109/1110/1111/1112/1113/1114/1115/1116/1117/1118/1119/1120/1121/1122/1123/1124/1125/1126/1127/1128/1129/1130/1131/1132/1133/1134/1135/1136/1137/1138/1139/1140/1141/1142/1143/1144/1145/1146/1147/1148/1149/1150/1151/1152/1153/1154/1155/1156/1157/1158/1159/1160/1161/1162/1163/1164/1165/1166/1167/1168/1169/1170/1171/1172/1173/1174/1175/1176/1177/1178/1179/1180/1181/1182/1183/1184/1185/1186/1187/1188/1189/1190/1191/1192/1193/1194/1195/1196/1197/1198/1199/1200/1201/1202/1203/1204/1205/1206/1207/1208/1209/1210/1211/1212/1213/1214/1215/1216/1217/1218/1219/1220/1221/1222/1223/1224/1225/1226/1227/1228/1229/1230/1231/1232/1233/1234/1235/1236/1237/1238/1239/1240/1241/1242/1243/1244/1245/1246/1247/1248/1249/1250/1251/1252/1253/1254/1255/1256/1257/1258/1259/1260/1261/1262/1263/1264/1265/1266/1267/1268/1269/1270/1271/1272/1273/1274/1275/1276/1277/1278/1279/1280/1281/1282/1283/1284/1285/1286/1287/1288/1289/1290/1291/1292/1293/1294/1295/1296/1297/1298/1299/1300/1301/1302/1303/1304/1305/1306/1307/1308/1309/1310/1311/1312/1313/1314/1315/1316/1317/1318/1319/1320/1321/1322/1323/1324/1325/1326/1327/1328/1329/1330/1331/1332/1333/1334/1335/1336/1337/1338/1339/1340/1341/1342/1343/1344/1345/1346/1347/1348/1349/1350/1351/1352/1353/1354/1355/1356/1357/1358/1359/1360/1361/1362/1363/1364/1365/1366/1367/1368/1369/1370/1371/1372/1373/1374/1375/1376/1377/1378/1379/1380/1381/1382/1383/1384/1385/1386/1387/1388/1389/1390/1391/1392/1393/1394/1395/1396/1397/1398/1399/1400/1401/1402/1403/1404/1405/1406/1407/1408/1409/1410/1411/1412/1413/1414/1415/1416/1417/1418/1419/1420/1421/1422/1423/1424/1425/1426/1427/1428/1429/1430/1431/1432/1433/1434/1435/1436/1437/1438/1439/1440/1441/1442/1443/1444/1445/1446/1447/1448/1449/1450/1451/1452/1453/1454/1455/1456/1457/1458/1459/1460/1461/1462/1463/1464/1465/1466/1467/1468/1469/1470/1471/1472/1473/1474/1475/1476/1477/1478/1479/1480/1481/1482/1483/1484/1485/1486/1487/1488/1489/1490/1491/1492/1493/1494/1495/1496/1497/1498/1499/1500/1501/1502/1503/1504/1505/1506/1507/1508/1509/1510/1511/1512/1513/1514/1515/1516/1517/1518/1519/1520/1521/1522/1523/1524/1525/1526/1527/1528/1529/1530/1531/1532/1533/1534/1535/1536/1537/1538/1539/1540/1541/1542/1543/1544/1545/1546/1547/1548/1549/1550/1551/1552/1553/1554/1555/1556/1557/1558/1559/1560/1561/1562/1563/1564/1565/1566/1567/1568/1569/1570/1571/1572/1573/1574/1575/1576/1577/1578/1579/1580/1581/1582/1583/1584/1585/1586/1587/1588/1589/1590/1591/1592/1593/1594/1595/1596/1597/1598/1599/1600/1601/1602/1603/1604/1605/1606/1607/1608/1609/1610/1611/1612/1613/1614/1615/1616/1617/1

124	ANOMI	I	ES	NT	ST	AC	I	Y	W	K	E	I	Y	A	S	C	I	T	I	ES	NC	IM	CC	T	A	D	Y	C	C	A	T	I	E	NC	CP	OT	183
-----	-------	---	----	----	----	----	---	---	---	---	---	---	---	---	---	---	---	---	---	----	----	----	----	---	---	---	---	---	---	---	---	---	---	----	----	----	-----

415 83NVCNDYTERECNUT VOMM VBNVA N DINVCNIM I 9003 T QMWC0004TMMY N 9 VZV3C101T1T MWNDDTS 4174

[illegible]

1960-1961

[illegible]

$\frac{1}{2}$

[illegible]

[illegible]

0 **1** **2** **3** **4** **5** **6** **7** **8** **9** **A** **B** **C** **D** **E** **F** **G** **H** **I** **J** **K** **L** **M** **N** **O** **P** **Q** **R** **S** **T** **U** **V** **W** **X** **Y** **Z**

fn

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

RESULT 23

C:Genetics:

A:Gene: pmpI

C:Superfamily: Chlamydomydia pneumoniae

Query Match 14.6%; Score 771.5; DB 2; Length 878;
Best Local Similarity 26.5%; Pred. No. 6.3e-39;
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

QY 53 DPSGTTVSAGELTUKNDLSIAALPLSCFQNL-----GSTVIGRGHSLTPE 101
DB 26 DPLGETA-----LTKNPNVCTPFEDCTMESLPALCAHQDDPLVILGNSCMFVS 80
QY 102 NIRTSTNGAALSNAADLFTIEGKELSPNCNSLLAVLPAATNKGSGQTPTTSTPSN 161
DB 81 KLHITDPEALFKEGD--LSIQNRFPLSFIDCSS-----KESSPS- 119
QY 162 GTIYSKT-DLLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVQENTAOADGAC 220
DB 120 -LIHQKNGQLSLRNGNSPCRNHABSGGAIADAPSLQHNLYLFTAFENSSKNGGAI 178
QY 221 QVITSFSAMANEAPIAFVNAVGVGGGIAAVQDQGGVSSSTSTEDPVPVFSRNVAEF 280
DB 179 QAQT-FSLSRVSPISFARNADLNGAIC----- 208
QY 281 DGNVARVGGIYSYGNVAPLNGKTLFLNNVASPVVIAAKQPTSGQASNTSNYDGGAI 340
DB 209 -----SNLICGNV-----NPLFF-----TGNATNGGAI 233
QY 341 FCKGAQAGSNNSVSDGEGVVPFSSNVAAGKGAIYAKKLSVANGPVQFLRNIAAD 400
DB 234 CCI--SDLNTSEKSLSLACQETLFASSAKGKAIYAKHVLRYNGPVFINNSAKI 291
QY 401 GGAIYLGSGHLSLADYDIIIDGNLKRKTAKENADVNGVTSQAISMGGKITTLLR 460
DB 292 GGAIAIGGSGLSILAGSGVLQNNQRTSQGLVR-NAIYLEKDAI-----LSSLE 343
QY 461 AKAGHQIILFNPI-EMANGNQPAQSSKLLKINDGEGTG-----DIVPA----- 504
DB 344 ARNG-DILFFPOIYQESSKESPLSSLOASVTSPTPATASPLVITQTSANKSVFSSRL 402
QY 505 -----NGSSTLYQNTYIGRVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
DB 403 SBEKTPDNLTSQLOQPIELKSGRLVKDRAVLAPSLSDQPALLIMEAGTS----- 455
QY 556 PPQPPAANQLITSLNHLSSLANNNAVTPNPPAQDSHPAVIGSTTAGSVTISGP 615
DB 456 -----LKTSDKLATLSIPLHSL-----DTEKSVTHAP 485
QY 616 -----IPFEDL-DDTAYDRYDMLGSNQKINVLKLGTKPPANAPSDLTL--GNEMPKY 666
DB 486 NLSIQKIFLSNGSDENFYENVLLSKQN-NIPLTL-----SKEQSHLHPDGNLSHP 539
QY 667 GYQGSWKLAMPNTANNPYTLKATWTGTGYNPGPVERVASLPNSLMGSIIDIRSAHSAI 726
DB 540 GYQGDWTFM--KDSDEG-HSLIANWTPKYVHPERQSTIVANTLWNTYSDMAQVSMI 596
QY 727 QASVDGRYCGLWVGVSNNFY-HDRDALG-QGYRTISGGYISGANSY-FCSSMFLGAF 783
DB 597 NTIAHGAGYLFPTWGSVSNLFYAHDSGKPTDNHHRSLGLYFGISHTSLDDHSFCLAA 656
QY 784 TEVGRSKDYVVCRSNHHACIGSVYLSYQALCGSYLFGDAPIASYCFGNQHKTSY-T 842
DB 657 GOLLGKSDSPITSTE-----TTSYIATVQALATPLM-KISAQACYNESITHEIKTKYRS 710
QY 843 FAEBSDVWDDNKLAGEIGAGLPIVITPSKLYLNLPRFPVQAEFSYADHESFTBEGDQAR 902
DB 711 FSKGFCGSHVAVSGEVSCASIPVNSGSLF-SFPSIFSKLQGFSGTQDQFESSGEIR 769
QY 903 AFKSHLLNLNVPGVKFDRCSSTHPNKYSFMAAYICDARTISGTETTLTSHQETWTD 962
DB 770 SFSASSFRNISLPMGITPEKKSQKTRNTYTFILGAYIQDLKRDVSGPVVLLKNAVSMADP 829
QY 963 AFLARHGVVVGSMYASLTNIEYVGHGREYDRDASRGYGLSAGSRVP 1012

Db 830 MANLDSRAYWRLTNQRAL-HRIQTLNVSIVLRGQSHSYSLDGLGTTTRF 878

RESULT 29

E86492

polymorphic outer membrane protein G family [imported] - Chlamydomydia pneumoniae (strain J138)
C:Species: Chlamydomydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: E86492
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise, T.; Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: E86492
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-712 <STO>
A:Cross-references: UNIPARC:UPI0000165697; GB:BA000008; NID:g8978386; PIDN:BA098223.1; G:G138
C:Experimental source: strain J138
C:Genetics:
A:Gene: pmp 2.1
C:Superfamily: Chlamydomydia pneumoniae polymorphic outer membrane protein G

Query Match 13.3%; Score 703; DB 2; Length 712;

Best Local Similarity 25.1%; Pred. No. 6.7e-35;

Matches 246; Conservative 109; Mismatches 304; Indels 320; Gaps 33;

QY 1 MQTSFHFKFFSLMILAYSCCSLNGGGYAAEIMVPGQIYDGETLTVSFPVTVIGDPSGTTVP 60
DB 1 MKIPLRLLISLVPFLSMNLLGAATTEELSASNS-FDGTTSSTSPSKTSSATDGTNYV 59
QY 61 SAGELTUKNDLSIAALPLSCFQNL--LLGSFTVIGRGHSLTPENI-RTSNGAALSAA 117
DB 60 FKDSVWIEVPKTGETQSTCFKNDAAAGDLNPLGGGFTFSFNIDATTASGAIGSEAA 119
QY 118 DGLFTIIGPKELSPNCNSLLAVLPAATNKGSGQTPTTSPNGTIYSKTDLLILNNEK 177
DB 120 NKTVTLSGFSALSP-----LKPASPTVTNGLGAINVKGNSLLDNDK 161
QY 178 PSFYSNLVSGDGAID-AKSLTVQGISKLCVFQENTAQADGACQVTVSFSAMANEAPIA 236
DB 162 VLIQDNFSTGCGAINCAGSLKI-----ANNKSL 191
QY 237 FVANAVGVGGIAAVQDQGGVSSSTSTEDPVPVFSRNVAEFDGPNVAVGGIYSYGN 296
DB 192 FIGNSSSTRGAIHT-----KN 208
QY 297 VAFLLNGKTLFLNNVASPVVIAAKQPTSGQASNTSNYDGGAIFCNGAAGSNNSGV 356
DB 209 LTLSSGGETLFGGNTA-----PT----- 226
QY 357 SFDGEGVVPFSSNVAAGKGAIYAKKLSVANGPVQFLRNIAANDGGAIVLGSGELSLA 416
DB 227 -----AAGK-----GGAIAADSTLSISG 246
QY 417 DYGDIIIPDGLKRTAKENADVNGVTSQAISMGGKITTIRAKAGHQLFPNDPIEMA 476
DB 247 DSGDIIPEG-----NTIGATG-TVSHSAIDLGTSAKITAKRAAQCHTIYFYDPTIV- 296
QY 477 NGNNQPAQSSKLLKIN-----DGEYTGDIVPA-----NGSSTLYQNTYIEQ 518
DB 297 TGSTSVADA---LNINSPTDGDNKBKTYGTIVFSGEKLTEAEKDKRRTSKLLQNVAFKN 353
QY 519 GRIVLRKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQPPQPPAANQLITSLNHLNLS 577
DB 354 GTVVVLKGDVLSANGFSQDANSKLMDLGTSL-----VANTSELTNLEINID 402
QY 578 SLLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDDTAYDR-----Y 629
DB 403 SLRNGKII-----KLSAATAQKDIRIDRPVVLAISDESFYQNGFLNEDHSY 448
QY 630 DMLGSNQKINVLKLGTKPPANAPSDTLGNENPKYGYQGSWKLAMPNTANNPYTLK 689

Db	449 D-----GILELDAGKDIVISADSRSDAVQSP-YGVQKWTINWSTDDK-----K 492	
Qy	690 AT--WTKGYNPGPERVASLVNSLWGSILDIRSAHSAIQASVDGRSVCRLGWGVGNF 747	
Db	493 ATVSNAKQSFNTABQEAFLPVNLLWGSFIDVRSFQNTIELGTEGAPYKRFWVAGISNV 552	
Qy	748 FPHDRDALCOGYRYISGGVSLGANSYP-GSSMFGLAFTVEVGRSKDYVVCRSNHHACIGS 806	
Db	553 LHRSGRENQKPRHVSGGAVGASTRMPGGDTLSLGAQLFARDKDYFNTNTFAKTYAGS 612	
Qy	807 VYLSQQALCGSYLFGDAFIRASVFGNQMKTSTYTFABESDVRWNNCLAGEIGAGLPI 866	
Db	613 LRL-----QHDASLYSVV-----SILLGE--GGLRE 636	
Qy	867 VTPSKLYNLELRP---FVQAEFSVADHESFTEE-----GDQARAFKSHL 909	
Db	637 ILLP---YVSKTLPSCFYQLSYGHTHRMKTESLPPPPRRRRRIILLGEDMSGLSEWEL 693	
Qy	910 -LNLSPVGVKDRCSSTH 927	
Db	694 ELLKIPAAEDFSK-STLH 711	
RESULT 30		
G72076		
polymorphic outer membrane protein h family - Chlamydophila pneumoniae (strain CWL029)		
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae		
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004		
C:Accession: G72076		
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, J.;		
Nature Genet. 21, 385-389, 1999		
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.		
A:Reference number: A72000; MUID:99206606; PMID:10192388		
A:Accession: G72076		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-978 <ARN>		
A:Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000046ABB; GB:AE001629; GB:AE001363; NID		
A:Experimental source: strain CWL029		
C:Genetics:		
A:Gene: pmp.14		
C:Superfamily: Polymorphic membrane protein H family		
Query Match 13.2%; Score 693.5; DB 2; Length 978;		
Best Local Similarity 25.6%; Pred. No. 4e-34;		
Matches 278; Conservative 147; Mismatches 412; Indels 251; Gaps 42;		
Qy	45 SPFYTVIGDPSGTVFVSAGELTLKNLD----NSIAALPLSCFGLNLAGSPFVLGRGHLTF 100	
Db	22 AFAETRLGGNFVPPITNQGBEILLTSDFVCSNFLGASFSSFTNSSNLSLLGKGLSLTF 81	
Qy	101 ENIRSTNGAALSNSAADGLFTIEGFKLSFSCNSLLAVLPAATTNKGSTQPTTTSTPS 160	
Db	82 TSCQAPTNYSYALLSAETL--TFKNPSSINFT-----GNQ-----STGL 119	
Qy	161 NGTIYKTDLLLNNEKFSYSLVSGDGAIDAKSLTVQISKLVCFOBNTQAQGGAC 220	
Db	120 GGLIYKG-DIVFQSIKDLIFTTNRV-----AYSPAVTTSATPAI-----TTVTTCASAL 168	
Qy	221 QVTVTSFMANEAPIAVANVAGVRGGIAAVDQGGQVSSSTSTEDPVVPSFNTA-VE 279	
Db	169 QPTDSLTVENISQISKPFGLNLFNPG-----SAISSSPFVAVVKFNTNTATMS 214	
Qy	280 FDGNVARVGGG-IYSYGNVAFPLNKGTLFLNNVASPVYIAAKQPTSQASNTSNVYDGG 338	
Db	215 FSNHTSSGGGVYGGSSLLFERNSSGCIIFTANSCVNSLKGVPSSSTYA-----LGSGG 269	
Qy	339 AIFCKNGAAGSNNSGVSFED-----GEGVFFSSNVAAKGKGA 378	
Db	270 AICPTGTPELKNQOGKCTSYNGTPNDAGAIYAETCNIVGNQALLDSDNTAARNGAI 329	
Qy	379 YAKKLSVANGCGVQFLNIANDGAIYLG-----ESGELSUSADYGIILFDGNLKRT 430	

Db	330 CAKVLNIQGRPIBFSRRAEKGGAIFGPSVGPAKQKTSTLTILASEGDIAFQGNMLNT 389	
Qy	431 AKENAADVNGVTVSSQAISMGSGKITTTRAKAGHOILFNDPIEMANGNNOPAOSSKLLK 490	
Db	390 KP-----GT-----RNAITVAGGEIVLSLQAQGSRLVFDPIYTHSLFTTSP--SNKDI 437	
Qy	491 INDGEGYTGDIVF-----ANGSSTLYONVTIEQGRIVLREKAKLSVNSLSQT 537	
Db	438 IN-ANGASGVVFTSKGLSSTELLIPANTTITLLGTVKIASGELKITDNVNVVLGFATQ 496	
Qy	538 G-GSLYMEAGSTWFTVTPPPQPPAANQLITLSNLHLSLSLLANNATVPTPTPPAQD 596	
Db	497 GSGQLTLGSGGTGLGLAT---PTGAPAAVD--FTIGKLAFPDFSFLEKRDVFS----- 542	
Qy	597 SHPAVIGSTTAG--SVTISGPIFPEDLDDTAYDRYDMLGNSQKINVLKQLQTKPEPANAP 654	
Db	543 -----ASVNAGTKNVTLTGALVLDEHVT--DLVD-----MWSLQSPVAIP 581	
Qy	655 SDLTILGNMPK-----YGYQGSWKLAWD-----PNTANNNGPYTLK 689	
Db	582 LAVFKGATVTKGPPDGEIATPSSHVGYQGKWSYTWSRPLLIAPDGGPFGGSPSANTLY 641	
Qy	690 ATW-----TKTGYNPGPERVASLVNSLWGSILDIRSAHSAIQ--ASVDGRSYCRLWYS 742	
Db	642 AVWNSDTLVRSTYILDPBRYGEIVSLSLWISFLGNQAFSDIILQDVLIDH-----PGLSIT 697	
Qy	743 GVSNEFFYHDDRDLG-----QGVRYISGGV-----SLGANSYFGSSMFLGAPTEVFOR 789	
Db	698 A-----KALGAYVHTPPQGHGSGRGGYQAALSMN-YTDHTYTLGLSFGQLYQK 747	
Qy	790 SKDYVVVCRSNHHACIGSVYLSLQQALCGSYLFG-----DAFI--RASYGFGNQHMK 838	
Db	748 TN-----ANPYDSRCSSEQMYLLS-----PFGQFPPIVTKSEALLISWKAAYGYSKNHLN 795	
Qy	839 TSY---TFABESDVRWNNCLAGETGAGLPI-----VTPSKLYLNELRPFVQAEFSYADH 891	
Db	796 TTYLRPDKAPKSGQGHNNSYVVLISAEHPFLNWCILTRPLAQAWDLDSGFISAEPLGGWQ 855	
Qy	892 ESFTREGDQARAFKSHLLNLSVPVGVKDFRCS-----THPNKYSFMAAYICDAYT 944	
Db	856 SKFTETGDLQRSFRSGKGYNVSLPIG-----CSSQWFTPFKKAPSTLTITKLAYKPDYIRV 910	
Qy	945 ISGTETTLSSHQRTWTTDAPHLARGVVRGSMYASLTSNIEVYGHGREYEDASRGYGL 1004	
Db	911 NPHNIVTVVSNQESISISGNLRRHGLFVQIHVDVLTEDTQAFNLNYTFDGKNGFTNHRV 970	
Qy	1005 SAGSRVRF 1012	
Db	971 STGLKSTF 978	
RESULT 31		
C86547		
polymorphic outer membrane protein H family [imported] - Chlamydophila pneumoniae (stra		
C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae		
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004		
C:Accession: C86547		
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I		
Nucleic Acids Res. 28, 2311-2314, 2000		
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.		
A:Reference number: A86491; MUID:20330349; PMID:10871362		
A:Accession: C86547		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-978 <STO>		
A:Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000131C83; GB:BA000008; NID:98978825; P		
A:Experimental source: strain J138		
C:Genetics:		
A:Gene: pmp.14		
C:Superfamily: Polymorphic membrane protein H family		
Query Match 13.1%; Score 688.5; DB 2; Length 978;		

Db	543	-----ASVNAKTWLTGALVLDHDT--DLVD-----MVSQSPVAIP	581
Qy	655	SDLTILGNMPK-----XYQGSWKLWD-----PNTANNQPYTLK	689
Db	582	IAVFKGATVTKTGFDPGBEIAATPSHYGYQKWSYTWSPRLIIPAPDGGPFGSPSANTILY	641
Qy	690	ATW-----TKTGNPQPPERVASLVNSLWGSILDIRSAHSAIQ--ASVDGRSYCYGLAWS	742
Db	642	AVWNSDTLVRSTYILDPERYGIIVNSLWISFLGNAQPSDILODVLILDH-----PGLSIT	697
Qy	743	GVSNPFPYHDDALG-----QGVYRISGGY-----SLGANSYFGSSMFLAFTVEVGR	789
Db	698	A-----KALGAYVHTPQGHGFGSGRYGQYQAALSMN-YTDHTTLGLSFGQLYCK	747
Qy	790	SKDYVYVCRSHHACIGSVYLSQOALCGSYLFG-----DAFI--RASVGFQGNQHK	838
Db	748	TN-----ANFYDRCSEOMYLLS-----FFGQPPIVTKSEALLISWKAAYGYSKHLN	795
Qy	839	TSY-----TFABESDVRWNNCLAGEICAGLPI-----VITPSKLYNLRLRPFVQAEPSYADH	891
Db	796	TYLRLDPKAPKQSQWHNSYYVLLISAEPFLNWCILLTRPLAQAWDLGFIAGFLGGWQ	855
Qy	892	ESTTEGDOARAPKSGHLNLSVPVGVKFDRCSS-----THPNKYSFMAAYICDAYRT	944
Db	856	SKFTETGDLQSPSRGKGNVSLPIG-----CSSQWTFPKKAPSTLTIKLAYKPDLYRV	910
Qy	945	ISGTETTLASHQSTTTDAPHLARHGVPVGRSVMYASLTSNIEVYGHGREYEDASRGYL	1004
Db	911	NPHNIVTVSNQSTISGANLRHGLFVQIHVDVLDLTDQAPLNTYTPDKNGFTNHRV	970
Qy	1005	SAGSRVP 1012	
Db	971	STGLKSTP 978	
RESULT 33			
F81722			
polymorphic membrane protein E/F family TC0261 [imported] - Chlamydia muridarum (strain C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn			
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004			
C;Accession: F81722			
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000			
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.			
A;Reference number: A81500; MUID:20150255; PMID:10684935			
A;Accession: F81722			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-976 <TET>			
A;Cross-references: UNIPROT:Q9PL47; UNIPARC:UPI0000057865; GB:AE002293; GB:AE002160; NID			
A;Experimental source: strain Nigg (MoPn)			
C;Genetics:			
A;Gene: TC0261			
Query Match 12.9%; Score 679.5; DB 2; Length 976;			
Best Local Similarity 26.1%; Pred. No. 2.9e-33;			
Matches 265; Conservative 150; Mismatches 413; Indels 187; Gaps 38;			
Qy	78	PLSCFGNLLGSFTVLGRHSIT---FENI-----RTSTNGAA-----LSNSAAD	118
Db	33	PYHMTGLFPFKVNLGLDTHNLTDYHLDNLKLCILACLOPTPYEGAAFTVDTYLSGSDTKD	92
Qy	119	GLPTIEGKELSPNCNSLLAVLPATNTKSGSQPTTSTPSNGTIYSKTDLLLNNEKF	178
Db	93	GIFC---FKNLT-----PESGGVIGSPQNTPIIKHNTI---GPVLFNNTCH	135
Qy	179	SFYSNL-----VSGDGAIDAKSLTVQGISKLCVFOENTAOGAGGCQVVT-----	224
Db	136	RLMTQTDPENEGNKAREGALHAGDVIYSNNQNLVGFKNFYVQGGAISANTPAYKENK	195
Qy	225	-SFSAMANEAPIAFVANYAVGRGGGIAAVQDQGGQSVSSTSTEDPVSFSRNTAVEFDGN	283

Db	196	SSFLCLANSC-----IQTKTGKGAAY-----VSTSCSFE-----NNNKDLFLFTQN	237
Qy	284	VARVCGGILYSYGNVAFLLNNGKTLFLNNVASPVYIAAKQPTSCQASNTSNNTYDGDGCAIFCK	343
Db	238	SGCAGGAIPTSCSLIGNQGDIVFVSNHGFK-----NVDNATNESDGGAI--K	284
Qy	344	NGAQAQSNNSGSVSDGEGVVFFPSNVAAGKGAIYAKKLKVANGCPVOFLNINIAN-DGG	402
Db	285	VTTRLDITNNGS-----QIFPSDNI SRNFGAIHAPCLHLVGNPTTYFTNNIANHTGG	337
Qy	403	AIYLGESGELSLSADYGDIIIDGNLKRKAKENAADVNGVTYSS-----QAISMGSGGKIT	457
Db	338	AIYITGTETSKISADHAIIPDNNI-----SANATNADGSSNTNPPHRAITWDSAGGI	393
Qy	458	TLRACAGHQILFNDPIEMANGN-----NOPAQSSKLLKINDGEGYTDGIVPANGSSTLYQ	512
Db	394	ELGAGKSONLIPYDPIQVNTAGVTDFNDASQTCVVFSGATVLSADISQANLQTKTPA	453
Qy	513	NVTIEQGRIVLRKAKLSVNSLSQTGSLYMEAG---STWDFVTTPQPOPPP---AANQL	566
Db	454	TLTSHGLLCIEDRAQLTVNNFTQTGGI VALGNGAVLSSYQHSITTDATQTPTTTTDDAS	513
Qy	567	ITLSNLHLSSLANNNAVTPNPTNPP--AOSHPAVIGSTTAGSVTISGPI--PFEDLD	622
Db	514	VTLNHIGLNLPSILKOGAEMPLLWVEPISTTQGNITTYTSDTAASFSLNGATLSLDEDG	573
Qy	623	DTAYDRYDLWLSNQKINVLKQLGTFK--PANAPSDLTIGNB-----MPKYGYQGSW	672
Db	574	NSPYENTD-----LSRALYAOPLMAISEASDNQLQSESMDFSKVNVPHYGMQGLW	623
Qy	673	KLAW-----DPNTANNQPYTLKATTKTCYNPQPPERVASLVNSLWGSIL	717
Db	624	TWGNAXTENPTTTPPATITDPKKAQNFHRTLLLTWLPACYIPSPKHSPLIANTLWGNIL	683
Qy	718	DIRS--AHSIAQASVDGRSYCRGLWVGSVNSFPYHDDRALGQGYRYSIGSYLG---ANS	772
Db	684	FATENLKNSSQGBELDRPFW--GITGGGLGMVYQSPRKHDPGFHMTSGYSAGMITGNT	741
Qy	773	YFGSSNFGLAFTVEVFCR-SKDYVYVCRSHHACIGSVYLSQOALCGSYLFGDAPIRASYG	831
Db	742	HTPSLRFSGSYTKLNERYAKNYV--SSKNYSQGEMLLSLQEGMLTKLIG-----LYS	793
Qy	832	FGNORHMTSYTFAEE--SDVRWNNCLAGEICAGLPI-----VITPSKLYNLRLRPFV	882
Db	794	YGNHNSHFFYTOGEDLSSQGEFHSQTFEGAVFDELPLKPPGRTHILT-----APPL	844
Qy	883	QAEFSYADHESPTTEGDOARAP-KSHLLNLSVPVGVKFDRCSSSTH-PNKYSFMAAYICD	940
Db	845	GAIGMYSKLSSPTEVGAYPRFTFETETPLINVLPIGVKGSFMNATHRPOAWTVELAYQPV	904
Qy	941	AYRTISGTETTLSSHQETWTTDAPHLARHGVPVGRSVMYASLT--SNIEVYGHGRY	993
Db	905	LYRQEPSSISTQLLAGKGMWFGHGPSASRALAYKISQKTQLLRFATLQOYHGTY	959
RESULT 34			
E71460			
probable outer membrane protein E - Chlamydia trachomatis (serotype D, strain UW3/Cx)			
C;Species: Chlamydia trachomatis			
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004			
C;Accession: E71460			
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell Science 282, 754-759, 1998			
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra			
A;Reference number: A71570; MUID:99000809; PMID:9784136			
A;Accession: E71460			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-964 <ARN>			
A;Cross-references: UNIPROT:O84877; UNIPARC:UPI00000131CF3; GB:AE001360; GB:AE001273; NI			
A;Experimental source: serotype D, strain UW-3/Cx			
C;Genetics:			


```
QY 727 QASVDGRSVCRLGWVGSVNPFFYHDDALG-QGYRIVISGGYSLGANSYFGSS-MFGLAFT 784
DB 656 SSWDEKHAASLOGIGL---LVHOKDKNGFGRFSGHMTGYSATTEATSSQSPNLSLGA 712
QY 785 YVFGRSKDYVV---CRSNHHACIGSVYLSLQALCGSYLFGDAFIRAS-----YGFQGNQHMK 838
DB 713 QPFSKAKEHESQNSSTSSH---YFS---GMCIENTLFPKHWRILSVSLAYMFTSEHTH 763
QY 839 TSYTFAER---SDVRDNNCLAGEIGAGLPVITPS-----KLYLNELRPPFQVQAFSYAD 890
DB 764 TMYQGLLEGNSQGSFHNHTLAG---ALSCVFLPQPHGESLQIY-----PFTALAIRGN 814
QY 891 HESFTEEGDQARAFKSGH---LLNLSPVGVKDRCSSTHPNKYSFM-----AAVICDAYRT 944
DB 815 LAAPQESGDHAREP-SLHRPLTDVSLPVGI---RASWKNHHRVPLVWLTEISYRSTLYRQ 870
QY 945 ISGTETTLTSHOBTWTTDAFHLARH---GVVVRGSM---YASLTSNIEVYGHGREYRDASR 1000
DB 871 DPFLSKLLISQGTWTTQATPVYNALGIKVKNTMQVFPKVTLSLDYSA----DISSSTL 926
QY 1001 GYLSAGSRVRF 1012
DB 927 SHYLNVASRMRF 938

RESULT 36
H72074
polymorphic membrane protein E/F family CP0286 [imported] - Chlamydomphila pneumoniae (sp
C:Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: H72074; E81593
R:Kalmann, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Pan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: H72074
A:Molecule type: DNA
A:Residues: 1-938 <ARN>
A:Cross-references: UNIPROT:Q92883; UNIPARC:UPI000004707C; GB:AE001631; GB:AE001363; NID
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: E81593
A:Molecule type: DNA
A:Residues: 1-938 <REA>
A:Cross-references: UNIPARC:UPI000004707C; GB:AE002190; GB:AE002161; NID:g7189209; PIDN:
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: pmp_15; CP0286

Query Match 12.2%; Score 644; DB 2; Length 938;
Best Local Similarity 25.8%; Pred. No. 3.9e-31;
Matches 266; Conservative 160; Mismatches 408; Indels 198; Gaps 44;

QY 81 CFCNGLG-SFTVLGRGHSITFENIRTSNGAALS---NSAAGLFTTEGPKELSP-SNCRN 135
DB 5 CFCNGLPFTFVLNAGLQPLEYIT-----LSPEYQAPQVGFTHNQDLAIVGNHN 58
QY 136 SLHVLPAATNKGSSQT-----PFTTSPNGTIYSKTDLLLNKEKSP 180
DB 59 DFILDYKYRSNGGALTCKNLLISENIGNVFPEKVCNPGSGAIYAAQNCTISKQNQYAF 118
QY 181 YSNLVSGD-----GGIDAASLTVOGSKLCVQENTAQADGACQVTSFSAMAN 231
DB 119 TTNLVSNPNTATAGSLGGALFAINCITSITNNLGGTFVDNLNKGALTYETNLSKDN 178
QY 232 EAPITAFVANVA---GVRRGGGIAAVQDQGGQGVSSSTSTEDPVSFSPRTAVEFDGNVARVG 288
DB 179 KGPITIKQNRALNSDSLGGGIYS-----GNSLNIE-----GNSGAIQITSSSGSG 224
```

```
QY 289 GGIYSYGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNYDGG-----AIFCKN 344
DB 225 GGIFSTQTLTISNNKLEISE-----NSAFANNYGSNPNFGGGLTTTFC-- 270
QY 345 GAQAGSNNSGSVSDGEGVFPSSNVAAGKGAIAKLSVANCGPVOPFLRNINDGGA 404
DB 271 -----TILNREGVL-FNNNQSQSNGGAIHAKSIIIKENGFPVFLNNTATRGAL 319
QY 405 YL-----GBSGLSLSADYDIIIFCNLKRXTAKENAAD--VNGVTVSSQAISGSGKITT 458
DB 320 LNLNAGSNGSGFILLSADNGDII FNNN---TASKHALNPPYRNAIHSTPNMNLQIG----- 371
QY 459 LRKAGHOILFNDPTIEMANGNNQPAQSKLLKINDGEGVTGDIIVFANGSSTLYQNVT--- 515
DB 372 --ARPGYVLFDPIE-----HELPSFPFILFNPETGHTGTVLFS--GEHVHONFTDEM 421
QY 516 -----IEQRIVLRKAKLSVNSLSQTGGSLYMEAGS--TWDPVTVPQPQPPA 562
DB 422 NPFYSYLRNTSELROGVLAVEDGAGLACYKFFQGGTLLLGQGAIVTTAGTIPTPSTPTT 481
QY 563 ANQLITLNLHLSLSLLANNAVTPNTPPAQDGHPAVIGST---TAGSVTISGPIFP 618
DB 482 VGSTITLNLHIALDLSILSFQA-----QAPKIWIYPTKGTSTYTEDSNPTITISGTLT 535
QY 619 EDL-DDTAYDRYDMLGSNQKINVLKL-QLGTPKPANAPSDLTGLNEMPKYIGQGSKLAW 676
DB 536 RNSNNEDPYDSLDLSHLSLEKVPVLLYVDVAAQKINSQOLDLSTLNSGEHYGYQGIWSTYW 595
QY 677 DNPT-----ANNPPTLKATWTKTYNPGPERVASLVPNSLWGSILDIRSAHSAI 726
DB 596 VETTTITNPTSLLGANTKHKLLYANWSPLYGRPHPERGEFITNALMQSAYTALAGLHSL 655
QY 727 QASVDGRSVCRLGWVGSVNPFFYHDDALG-QGYRIVISGGYSLGANSYFGSS-MFGLAFT 784
DB 656 SSWDEKHAASLOGIGL---LVHOKDKNGFGRFSGHMTGYSATTEATSSQSPNLSLGA 712
QY 785 YVFGRSKDYVV---CRSNHHACIGSVYLSLQALCGSYLFGDAFIRAS-----YGFQGNQHMK 838
DB 713 QPFSKAKEHESQNSSTSSH---YFS---GMCIENTLFPKHWRILSVSLAYMFTSEHTH 763
QY 839 TSYTFAER---SDVRDNNCLAGEIGAGLPVITPS-----KLYLNELRPPFQVQAFSYAD 890
DB 764 TMYQGLLEGNSQGSFHNHTLAG---ALSCVFLPQPHGESLQIY-----PFTALAIRGN 814
QY 891 HESFTEEGDQARAFKSGH---LLNLSPVGVKDRCSSTHPNKYSFM-----AAVICDAYRT 944
DB 815 LAAPQESGDHAREP-SLHRPLTDVSLPVGI---RASWKNHHRVPLVWLTEISYRSTLYRQ 870
QY 945 ISGTETTLTSHOBTWTTDAFHLARH---GVVVRGSM---YASLTSNIEVYGHGREYRDASR 1000
DB 871 DPFLSKLLISQGTWTTQATPVYNALGIKVKNTMQVFPKVTLSLDYSA----DISSSTL 926
QY 1001 GYLSAGSRVRF 1012
DB 927 SHYLNVASRMRF 938

RESULT 37
H71460
probable outer membrane protein H - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
R:Accession: H71460
R:Stephens, R.S.; Kalmann, S.; Lammel, C.J.; Pan, J.; Marathe, R.; Mitchell
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: H71460
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1016 <ARN>
A:Cross-references: UNIPROT:O84880; UNIPARC:UPI0000046AC7; GB:AE001360; GB:AE001273; NID
A:Experimental source: serotype D, strain UW-3/Cx
```


Db 413 PNPEHLGTLTILFSSAVIPDTSTSRDDPFISHRNHIGLYNGTLEALDRAEKVKYKFPQFGGT 472
Qy 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTN---PPTN---PPA 594
Db 473 LRL--GSRVAVFTTDEQSSSVGVSVININLAINLPSILGNRVAPKLRWIRTPGSSAPYS 530
Qy 595 QDSHPAVIGSTTAGSVTISGPIFFEDLDATAYDRYDMLGSNQ---KINVLKQLQGTGKPPA 651
Db 531 EDNNPI-----INLSGPLSL--LDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHI 579
Qy 652 NAPSDLTLG-NEMPKYGYGSKWLAW-----DPNTANGPYTLKATTKTGYNPG 700
Db 580 NTDNFPYEGTLNTQHYGYGVNSPYWIEITITSDTSSEDTVNTLHRLQYGDWTPGTGKYN 639
Qy 701 PERVASLVPNSLWGS-----LDIRSAHSAQASVDGRSYCRGLWVSGVSNFFYH---D 751
Db 640 PENKGDIALSAFWQSFHNLFATLRYTQOQOIAPTASGEA-----TRLFVHQN 689
Qy 752 RDALGGYRYISGGYSLGANSYFGSS-MFGLAFTEVF-----GRSKDYVVCRRNNHAC-I 804
Db 690 NDA--KGFMEATGYSIGTNTASNHSFGVNSQLFSLNLYESHSDNSVASHTTTVALQI 747
Qy 805 GSVLTSTQALCGSYLFGDAFIRASYGFGNQHKTSYTFABESDVRWNNCLAGETGAGL 864
Db 748 NNPWLQERFSTAS-----LAYSNSHHIKAS---GYSKIQTEGKCYSTTLGAAL 795
Qy 865 -----PIVITPSKLYLNELPFVQAEPFQADHESFTTEGQOARAFKSGH--LLN 911
Db 796 SCLSLQWRSRPLHFT-----PFIQAIARNSQTAFOESGDKARKE-SVHKPLYN 844
Qy 912 LSPVGV-----KPDRCSTHNPKNYSFMAAYICDAYRTISGTTETTLTSHOETWTTDAFH 965
Db 845 LTVPLGIQSAWESKF-----RLPYWNIELAYQPVLYQQNPVNSLESSGSSWLLSGIT 899
Qy 966 LARHGVMVVRGSMYASLTNSIEVYGHGREYDRASRGYGLSAGSRVRF 1012
Db 900 LARNAIAPKGRNQIFIPPKLSVFLDYQGSVSSSTTTTHLHAGTTKFK 946

RESULT 39
D81594
polymorphic membrane protein E/F family CP0283 [imported] - Chlamydia pneumoniae (se
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81594
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: AB1500; MUID:20150255; PMID:10684935
A;Accession: D81594
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-946 <REA>
A;Cross-references: UNIPROT:Q92880; UNIPARC:UPI00000131C86; GB:AE002161; NII
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0283

Query Match 11.6%; Score 609; DB 2; Length 946;
Best Local Similarity 24.7%; Pred. No. 5.2e-29;
Matches 264; Conservative 138; Mismatches 419; Indels 246; Gaps 43;

Qy 60 PSAGELTKNLDNSIAALPLS-CFGLNLSGFTVLGRGHSITPENIRSTNGAALSNAAD 118
Db 12 FPGCALILGKTTILNATPLSDYFDNQANQLTTL-----FPLIDLTNMTPTYSRA-- 62
Qy 119 GLFTIEGFKELSFNCNSLLAVLPAATNKGSGQPTTSTPSNGTIYSKTDLLILNNEKF 178
Db 63 ---TLFGVRD-----DTNQ-----DIVLDHONSI 83
Qy 179 SPYSLVSGDGAIDAKSLTYQGISKLCVQFNTAQADGGACQVVTSPSAMANEAPAFV 238

Db 84 ESWFENFSQDGGALSCSKSLAITNTNQNLPLNSPAIKRAGAMYVNGNFDLSENHGSILFS 143
Qy 239 ANVAGVRGGGIAAVODGQGVSSSTSTEDPVVVSFRNTAVBPDGNVAVRGGIY----- 292
Db 144 GNLSPFNASNADTCTCGAVLCSKNVT---ISKQGTAY-FINKKAKSSGGAIQAAIIN 198
Qy 293 ---SYGNVAFPLN-----GKTLFLN-----NVASPVVIAAKOPTSGQA----- 327
Db 199 IKDNTGPCLFFNNAAGTAGGALFANACRIENNSQPIYPLNNQSGLGGAIRVHOCILTK 258
Qy 328 -----SNTSNNTYGDGAIFCGNQAAGSNNNSGVSPDGEVGVFSSNVAAG 373
Db 259 NTGSVIFNNFAMEADIIANHSGGAIYC-----ISCSIK-DNPGIAAFDNNNTAAR 308
Qy 374 KGAIAVAKKLSVANQGPVQFLNIANDGGAIVLGSBELSLSDADYDIIIFDGNLKRKTAKE 433
Db 309 DGAICTQSLTIQDSGPVFTNNQGTGGAIMLRDGACTLFDADQGDIIIFYNN--RHFKD 366
Qy 434 NAADVNGVTVS--SQAISMGSGGKIITLRKAGHQILFNDPIEMANGNNOQPAOSSKLLKIN 492
Db 367 TPS--NHVSVNCTRNVS-----TVGASQGHSAFTFYDPIL-----QRYTIQNS-IOKFN 412
Qy 493 DGEYGTGDIVFANG-----SSTLYQNVTIEQGRIVLREKAKLSVNSLSOTGGS 540
Db 413 PNPEHLGTLTILFSSAVIPDTSTSRDDPFISHRNHIGLYNGTLEALDRAEKVKYKFPQFGGT 472
Qy 541 LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLSLSLLANNVNTN---PPTN---PPA 594
Db 473 LRL--GSRVAVFTTDEQSSSVGVSVININLAINLPSILGNRVAPKLRWIRTPGSSAPYS 530
Qy 595 QDSHPAVIGSTTAGSVTISGPIFFEDLDATAYDRYDMLGSNQ---KINVLKQLQGTGKPPA 651
Db 531 EDNNPI-----INLSGPLSL--LDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHI 579
Qy 652 NAPSDLTLG-NEMPKYGYGSKWLAW-----DPNTANGPYTLKATTKTGYNPG 700
Db 580 NTDNFPYEGTLNTQHYGYGVNSPYWIEITITSDTSSEDTVNTLHRLQYGDWTPGTGKYN 639
Qy 701 PERVASLVPNSLWGS-----LDIRSAHSAQASVDGRSYCRGLWVSGVSNFFYH---D 751
Db 640 PENKGDIALSAFWQSFHNLFATLRYTQOQOIAPTASGEA-----TRLFVHQN 689
Qy 752 RDALGGYRYISGGYSLGANSYFGSS-MFGLAFTEVF-----GRSKDYVVCRRNNHAC-I 804
Db 690 NDA--KGFMEATGYSIGTNTASNHSFGVNSQLFSLNLYESHSDNSVASHTTTVALQI 747
Qy 805 GSVLTSTQALCGSYLFGDAFIRASYGFGNQHKTSYTFABESDVRWNNCLAGETGAGL 864
Db 748 NNPWLQERFSTAS-----LAYSNSHHIKAS---GYSKIQTEGKCYSTTLGAAL 795
Qy 865 -----PIVITPSKLYLNELPFVQAEPFQADHESFTTEGQOARAFKSGH--LLN 911
Db 796 SCLSLQWRSRPLHFT-----PFIQAIARNSQTAFOESGDKARKE-SVHKPLYN 844
Qy 912 LSPVGV-----KPDRCSTHNPKNYSFMAAYICDAYRTISGTTETTLTSHOETWTTDAFH 965
Db 845 LTVPLGIQSAWESKF-----RLPYWNIELAYQPVLYQQNPVNSLESSGSSWLLSGIT 899
Qy 966 LARHGVMVVRGSMYASLTNSIEVYGHGREYDRASRGYGLSAGSRVRF 1012
Db 900 LARNAIAPKGRNQIFIPPKLSVFLDYQGSVSSSTTTTHLHAGTTKFK 946

RESULT 40

C72075
polymorphic outer membrane protein e/f family - Chlamydia pneumoniae (strain CWL029
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: C72075
R;Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2006, 12:15:19 ; Search time 243 Seconds
(without alignments)
2938.250 Million cell updates/sec

Title: US-10-701-844-2

Perfect score: 5267

Sequence: 1 MQTSPHKFSLMILAYSCS.....YETDASRGYGLSAGSRVRF 1012

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5131.5	97.4	1013	1	PMPG_CHLTR
2	3845	73.0	987	1	PMPG_CHLMU
3	1860.5	35.3	1011	2	Q823X5_CHLCV
4	1776	33.7	1024	2	Q5L6J7_CHLAB
5	1365.5	25.9	973	1	PMP13_CHLPN
6	1158	22.0	942	2	Q823X1_CHLCV
7	1122	21.3	926	2	P71135_CHLAB
8	1122	21.3	926	2	Q5L6J3_CHLAB
9	1101	20.9	928	1	PMP10_CHLPN
10	1099	20.9	926	2	Q823W9_CHLCV
11	1079	20.7	928	1	PMP11_CHLPN
12	1079	20.5	930	1	PMP8_CHLPN
13	1077	20.4	1378	2	Q5L6J2_CHLAB
14	1040	19.7	936	1	PMP7_CHLPN
15	1021	19.4	928	1	PMP9_CHLPN
16	1012	19.2	1276	1	PMP6_CHLPN
17	1008	19.1	866	2	Q823X0_CHLCV
18	978.5	18.6	772	2	Q9RB71_CHLPN
19	971.5	18.4	841	2	Q822Q5_CHLCV
20	967	18.4	922	1	PMP1_CHLPN
21	950	18.0	839	2	P71132_CHLAB
22	949.5	18.0	839	2	P77792_CHLAB
23	939.5	17.8	841	1	PMP2_CHLPN
24	897	17.0	846	2	P71133_CHLAB
25	803	15.2	843	2	Q823X2_CHLCV
26	789	15.0	867	1	PMP1_CHLMU
27	787.5	15.0	602	2	Q8VU49_CHLPS
28	785.5	14.9	843	2	Q823X3_CHLCV
29	785	14.9	849	2	Q5L6J5_CHLAB
30	784.5	14.9	878	2	Q4VWS6_CHLTR
31	781.5	14.8	878	2	Q4VWR2_CHLTR

32	779.5	14.8	878	2	Q4VWS3_CHLTR	Q4VWS3 chlamydia t
33	779.5	14.8	878	2	Q4VWS8_CHLTR	Q4VWS8 chlamydia t
34	777.5	14.8	878	2	Q4VWR8_CHLTR	Q4VWR8 chlamydia t
35	773.5	14.7	840	2	Q5L6J6_CHLAB	Q5L6J6 chlamydophi
36	771.5	14.6	846	2	Q84FU4_CHLTR	Q84FU4 chlamydia t
37	771.5	14.6	878	1	PMP1_CHLTR	Q84882 chlamydia t
38	771.5	14.6	878	2	Q4VWS2_CHLTR	Q4VWS2 chlamydia t
39	770.5	14.6	878	2	Q4VWR1_CHLTR	Q4VWR1 chlamydia t
40	769.5	14.6	846	2	Q84FU2_CHLTR	Q84FU2 chlamydia t
41	769.5	14.6	878	2	Q4VWR6_CHLTR	Q4VWR6 chlamydia t
42	766.5	14.6	846	2	Q84FU1_CHLTR	Q84FU1 chlamydia t
43	766.5	14.6	846	2	Q84FU3_CHLTR	Q84FU3 chlamydia t
44	764.5	14.5	846	2	Q84FU0_CHLTR	Q84FU0 chlamydia t
45	764.5	14.5	846	2	Q84FU5_CHLTR	Q84FU5 chlamydia t

ALIGNMENTS

RESULT 1

PMPG_CHLTR
ID_PMPG_CHLTR STANDARD; PRT; 1013 AA.
AC O84879;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G).
DE protein G).
GN Name=pmpG; OrderedLocusNames=CT871;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=D/UM-3/Cx;
RX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.
RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis."
RT Science 282:754-759(1998).
RL -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AR001360; AAC69469.1; -; Genomic_DNA.
PIR; G71460; G71460.

PHCI-2DPAGE; O84879; -.

InterPro; IPR005546; Auto transptbeta.

InterPro; IPR006315; Autotransporter.

InterPro; IPR011427; ChlamPMP_M.

InterPro; IPR003368; Chlamydia_PMP.

Pfam; PF03797; Autotransporter; 1.

Pfam; PF02415; Chlam_PMP; 9.

Pfam; PF07548; ChlamPMP_M; 1.

TIGRFAM; TIGR01414; autotrans_bar1; 1.

TIGRFAM; TIGR01376; POMP repeat; 5.

Complete proteome; Membrane; Multigene family; Outer membrane; Signal.

SIGNAL 1 27 Potential.

CHAIN 28 1013 Probable outer membrane protein pmpG.

SEQUENCE 1013 AA; 107367 MW; F0927743C0A651DD CRC64;

Query Match 97.4%; Score 5131.5; DB 1; Length 1013;

Best Local Similarity 97.0%; Pred. NO. 3.1e-295;

Matches 993; Conservative 18; Mismatches 11; Indels 1; Gaps 1;


```

QY 1 MQTSPHKKFLSHLAYSCCSLNGGCGYAAEIMVPGIYDGETLTIVSPFYTVIGDPSGTTVP 60
DB 1 MQTSPHKKFLSHLAYSCCSLNGGCGYAAEIMVPGIYDGETLTIVSPFYTVIGDPSGTTVP 60
QY 61 SAGELTLKLNLDNSIAALPLSCFNLGSLTFTVLRGHSITFENIRTSNGAALSNSAADGL 120
DB 61 SAGELTLKLNLDNSIAALPLSCFNLGSLTFTVLRGHSITFENIRTSNGAALSNSAADGL 120
QY 121 FTIEGFKELSPNCNLSLAVLPAAATNNKSGSTPTTSTPSNGTIYSKTDLLLNNKESFP 180
DB 121 FTIEGFKELSPNCNLSLAVLPAAATNNKSGSTPTTSTPSNGTIYSKTDLLLNNKESFP 180
QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVQENHTAQADGGACQVVTSSAMANEAPITAFVN 240
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVQENHTAQADGGACQVVTSSAMANEAPITAFVN 240
QY 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEPDGNVARVGGGIYSYGNVAPL 300
DB 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEPDGNVARVGGGIYSYGNVAPL 300
QY 301 NNGKTLFLNNVSPVYIAAKQPTSGASNTSNNGYDGGAICTONGAQ-AGSNNSGSVSD 359
DB 301 NNGKTLFLNNVSPVYIAAKQPTSGASNTSNNGYDGGAICTONGAQ-AGSNNSGSVSD 359
QY 360 GEGVFPSSNVAAGKGAIYAKLSVANGCPVQFLRNTANDGGAIYIGRSGLSLSDYG 419
DB 360 GEGVFPSSNVAAGKGAIYAKLSVANGCPVQFLRNTANDGGAIYIGRSGLSLSDYG 419
QY 420 DIIPDGNLKRTAKENAAVNGVTVSSQAI SMGSGKITTLAKAGHQLTLPNDPIEMANGN 479
DB 420 DIIPDGNLKRTAKENAAVNGVTVSSQAI SMGSGKITTLAKAGHQLTLPNDPIEMANGN 479
QY 480 NPAQSSKLLKINDGEGTIDIVFANGSSTLYQNTIIEQGRIVLREKAKLSVNSLSQTGG 539
DB 480 NPAQSSKLLKINDGEGTIDIVFANGSSTLYQNTIIEQGRIVLREKAKLSVNSLSQTGG 539
QY 540 SLYMEAGTDFVTPPQQPQAPPAANQLITLNLHLSLSLANNAVTPNPQADSHP 599
DB 540 SLYMEAGTDFVTPPQQPQAPPAANQLITLNLHLSLSLANNAVTPNPQADSHP 599
QY 600 AVIGSTTAGSVTISGPIEFEDLDDTAYDRYDNLGSKNIVLKLQGTGKPPANAPSDTL 659
DB 600 AVIGSTTAGSVTISGPIEFEDLDDTAYDRYDNLGSKNIVLKLQGTGKPPANAPSDTL 659
QY 660 QNEMPKYQGSWKLAMPDNTANNPGLTKATWTGYNPGPERSVASLVPNSLWGSILDI 719
DB 660 QNEMPKYQGSWKLAMPDNTANNPGLTKATWTGYNPGPERSVASLVPNSLWGSILDI 719
QY 720 RSHASIAQASVDRGSRVCRGLWVGSVSNFPHYDRDALQGYRYISGGYISLGANSYFGSMP 779
DB 720 RSHASIAQASVDRGSRVCRGLWVGSVSNFPHYDRDALQGYRYISGGYISLGANSYFGSMP 779
QY 780 GLAPTEVFGSKDYVVCSSNHACIGSVYLSLQALCGSLYFGDAPIRASVYGRGNHMT 839
DB 780 GLAPTEVFGSKDYVVCSSNHACIGSVYLSLQALCGSLYFGDAPIRASVYGRGNHMT 839
QY 840 SYTFAESDVRWDDNNCLAGEIAGLPIVITPSKLYLNEIRLPVQAEFSDYADHESFTEGD 899
DB 840 SYTFAESDVRWDDNNCLAGEIAGLPIVITPSKLYLNEIRLPVQAEFSDYADHESFTEGD 899
QY 900 QARAFKSGHLNLVSVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959
DB 900 QARAFKSGHLNLVSVGVKFDRCSSSTHPNKYSFMAAYICDAYRTISGTETLLSHQETW 959
QY 960 TTDAPHLARHGIVVGRGSMVYASITSNIEVYGRVYRDASRGYGLSAGSVRFP 1012
DB 960 TTDAPHLARHGIVVGRGSMVYASITSNIEVYGRVYRDASRGYGLSAGSVRFP 1012

```

RESULT 2

PMPG_CHLWU

ID_PMPG_CHLWU

STANDARD;

PRT;

987 AA.

```

AC 09PL45;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane
DB protein G).
GN Name-pmpG; OrderedLocusNames=TC0263;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MoPn / Nig99;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Uetback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson M.C., DeBoy R.T., Kolonay J.P.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.; Chlamydia
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -|- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -|- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -|- SIMILARITY: Contains 1 autotransporter (TC 1.9.12) domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB002293; AAF39132.1; -; Genomic_DNA.
DR PIR; H81722; H81722.
DR TIGR; TC0263; -.
DR InterPro; IPR005546; Auto_transp_beta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 9.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 5.
DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
KW SIGNAL
FT CHAIN 26 987 Probable outer membrane protein pmpG.
FT SEQUENCE 987 AA; 104867 MW; 32079BD6BEB2DA42 CRC64;
QY 1 MQTSPHKKFLSHLAYSCCSLNGGCGYAAEIMVPGIYDGETLTIVSPFYTVIGDPSGTTVP 60
DB 1 MQTSPHKKFLSHLAYSCCSLNGGCGYAAEIMVPGIYDGETLTIVSPFYTVIGDPSGTTVP 60
QY 61 SAGELTLKLNLDNSIAALPLSCFNLGSLTFTVLRGHSITFENIRTSNGAALSNSAADGL 120
DB 61 SAGELTLKLNLDNSIAALPLSCFNLGSLTFTVLRGHSITFENIRTSNGAALSNSAADGL 120
QY 121 FTIEGFKELSPNCNLSLAVLPAAATNNKSGSTPTTSTPSNGTIYSKTDLLLNNKESFP 179
DB 121 FTIEGFKELSPNCNLSLAVLPAAATNNKSGSTPTTSTPSNGTIYSKTDLLLNNKESFP 179
QY 181 YSNLVSGDGAIDAKSLTVQGISKLCVQENHTAQADGGACQVVTSSAMANEAPITAFVN 239
DB 181 YSNLVSGDGAIDAKSLTVQGISKLCVQENHTAQADGGACQVVTSSAMANEAPITAFVN 239
QY 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEPDGNVARVGGGIYSYGNVAPL 299
DB 241 VAGVRGGGIAAVQDQGGVSSSTSTEDPVVFSRNTAVEPDGNVARVGGGIYSYGNVAPL 299
QY 231 NVAGNKGGVAAVVDGQ-QGAGGAT---DLSVNPANNTAVEFGNSARIIGGDIYDGNISFP 286
DB 231 NVAGNKGGVAAVVDGQ-QGAGGAT---DLSVNPANNTAVEFGNSARIIGGDIYDGNISFP 286

```

Query Match 73.0%; Score 3845; DB 1; Length 987;
 Best Local Similarity 72.6%; Pred. No. 5.1e-219;
 Matches 739; Conservative 105; Mismatches 136; Indels 38; Gaps 13;

300 LNNKTLFLNNVSPVYI--AAKQPTSGQASNTSNNGYDGGGAIKCKNGAAGSNNSSGVS 357
 DB LNAATKTLFLNNVSPVYI--AAKQPTSGQASNTSNNGYDGGGAIKCKNGAAGSNNSSGVS 358
 358 FDGEGVFPSSNAAGKGAIAKGLSVANCGPQVFLRNANDGAIYLGESGELSLSAD 417
 DB FDGEGVFPSSNAAGKGAIAKGLSVANCGPQVFLRNANDGAIYLGESGELSLSAD 418
 339 FDGEGVFPSSNAAGKGAIAKGLSVANCGPQVFLRNANDGAIYLGESGELSLSAD 398
 DB FDGEGVFPSSNAAGKGAIAKGLSVANCGPQVFLRNANDGAIYLGESGELSLSAD 399
 418 YGDIIPDGNLKTAKENADVNGVTVSSQAIKNGSGGKITTILRAKAGHQLFNDPIEMAN 477
 DB YGDIIPDGNLKTAKENADVNGVTVSSQAIKNGSGGKITTILRAKAGHQLFNDPIEMAN 478
 399 RGDIIIFDGNLKTAKENADVNGVTVSSQAIKNGSGGKITTILRAKAGHQLFNDPIEMAN 458
 DB RGDIIIFDGNLKTAKENADVNGVTVSSQAIKNGSGGKITTILRAKAGHQLFNDPIEMAN 459
 478 GNNQPAQSSKLLKINDGEGYTDIVPANGSSTLYQNVTIEQGRIVILREKAKLSVNSLSQT 537
 DB GNNQPAQSSKLLKINDGEGYTDIVPANGSSTLYQNVTIEQGRIVILREKAKLSVNSLSQT 538
 459 G--QPV--TQTLTVNEGEGYTDIVPANGSSTLYQNVTIEQGRIVILREKAKLSVNSLSQT 514
 DB G--QPV--TQTLTVNEGEGYTDIVPANGSSTLYQNVTIEQGRIVILREKAKLSVNSLSQT 515
 538 GGSGLVMEAGSTWDFVTPQPPQPPAANQLITLSNHLHLSLLANNAVTPNTPPAQDS 597
 DB GGSGLVMEAGSTWDFVTPQPPQPPAANQLITLSNHLHLSLLANNAVTPNTPPAQDS 598
 515 GGSVMEAGSTWDFVTPQPPQPPAANQLITLSNHLHLSLLANNAVTPNTPPAQDS 569
 DB GGSVMEAGSTWDFVTPQPPQPPAANQLITLSNHLHLSLLANNAVTPNTPPAQDS 570
 598 HPAVIGSTAGVTTISGPIFFEDLDOTAYDRYDMLGSKNOKINVLKQLGTGKPPANAPSDL 657
 DB HPAVIGSTAGVTTISGPIFFEDLDOTAYDRYDMLGSKNOKINVLKQLGTGKPPANAPSDL 658
 570 SPAVIGNTAAGTVTTISGPIFFEDLDOTAYDRYDMLGSKNOKINVLKQLGTGKPPANAPSDL 629
 DB SPAVIGNTAAGTVTTISGPIFFEDLDOTAYDRYDMLGSKNOKINVLKQLGTGKPPANAPSDL 630
 658 TGNEMPKYGYOGSKWLAWDPNTAN---NGPYTLKATWTKGVNPGPERVASI.VPNSLWG 714
 DB TGNEMPKYGYOGSKWLAWDPNTAN---NGPYTLKATWTKGVNPGPERVASI.VPNSLWG 715
 630 TGNEMPKYGYOGSKWLAWDPNTAN---NGPYTLKATWTKGVNPGPERVASI.VPNSLWG 689
 DB TGNEMPKYGYOGSKWLAWDPNTAN---NGPYTLKATWTKGVNPGPERVASI.VPNSLWG 690
 715 SILDTRSAHSAQASVDGSHYCRGLWVSGVSPFFHYDRDALGQGYRISYGSYSLGANSYF 774
 DB SILDTRSAHSAQASVDGSHYCRGLWVSGVSPFFHYDRDALGQGYRISYGSYSLGANSYF 775
 690 SILDTRSAHSAQASVDGSHYCRGLWVSGVSPFFHYDRDALGQGYRISYGSYSLGANSYF 749
 DB SILDTRSAHSAQASVDGSHYCRGLWVSGVSPFFHYDRDALGQGYRISYGSYSLGANSYF 750
 775 GSSMFLAPTEVGRSKDYVCRSHHACIGSVYLSQOALCGSYLFGDAPTRASYGFGN 834
 DB GSSMFLAPTEVGRSKDYVCRSHHACIGSVYLSQOALCGSYLFGDAPTRASYGFGN 835
 750 GSSMFLAPTEVGRSKDYVCRSHHACIGSVYLSQOALCGSYLFGDAPTRASYGFGN 809
 DB GSSMFLAPTEVGRSKDYVCRSHHACIGSVYLSQOALCGSYLFGDAPTRASYGFGN 810
 835 QHMKTSYTTAEBSVDVNDNCLAGEIGAGLPIVITPSKLYLNLRLPVOAEPSYADHESP 894
 DB QHMKTSYTTAEBSVDVNDNCLAGEIGAGLPIVITPSKLYLNLRLPVOAEPSYADHESP 895
 810 QHMKTSYTTAEBSVDVNDNCLAGEIGAGLPIVITPSKLYLNLRLPVOAEPSYADHESP 869
 DB QHMKTSYTTAEBSVDVNDNCLAGEIGAGLPIVITPSKLYLNLRLPVOAEPSYADHESP 870
 895 TREGDQARAFKSHLNLVSPVGVKFDRCSTHNPKNYSFMAAYICDAYRTISGTETTLAS 954
 DB TREGDQARAFKSHLNLVSPVGVKFDRCSTHNPKNYSFMAAYICDAYRTISGTETTLAS 955
 870 TREGDQARAFKSHLNLVSPVGVKFDRCSTHNPKNYSFMAAYICDAYRTISGTETTLAS 929
 DB TREGDQARAFKSHLNLVSPVGVKFDRCSTHNPKNYSFMAAYICDAYRTISGTETTLAS 929
 955 HQTETTTDAPHLARHGVMVGRGMYASLTNIEYVGHRYEYRDAARGYGLSAGSRVRP 1012
 DB HQTETTTDAPHLARHGVMVGRGMYASLTNIEYVGHRYEYRDAARGYGLSAGSRVRP 987
 930 HQTETTTDAPHLARHGVMVGRGMYASLTNIEYVGHRYEYRDAARGYGLSAGSRVRP 987
 DB HQTETTTDAPHLARHGVMVGRGMYASLTNIEYVGHRYEYRDAARGYGLSAGSRVRP 987

RESULT 3
 Q823X5 CHLCV PRELIMINARY; PRT; 1011 AA.
 AC Q823X5 CHLCV PRELIMINARY; PRT; 1011 AA.
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Polymorphic outer membrane protein G family
 DE protein/autotransporter.
 DE OrderedLocusNames=CCA00278;
 GN Chlamydia caviae.
 OS Chlamydia caviae.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83557;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GPIC;
 RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
 RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
 RA Heidelberg J.F., Holtzapfel B.K., Khouri H.M., Federova N.B.,
 RA Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
 RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
 RA Bavoil P.M., Fraser C.M.;
 RA "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):

examining the role of niche-specific genes in the evolution of the
 Chlamydiaceae.";
 RT Nucleic Acids Res. 31:2134-2147 (2003).
 RL EMBL; AB016995; AAP05029.1; -, Genomic_DNA.
 DR TIGR; CCA00278;
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Auto transportbeta.
 DR InterPro; IPR011427; Chlamydia_PMP.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF07448; Chlamydia_PMP; 1.
 DR Pfam; PF02415; Chlamydia_PMP; 7.
 DR TIGRPFAMs; TIGR01414; autotrans bar1; 1.
 DR TIGRPFAMs; TIGR01376; POMP_repeat; 6.
 KW Complete proteome.
 SQ SEQUENCE 1011 AA; 107952 MW; EBD9A9FALC46EDBE6 CRC64;
 Query Match 35.3%; Score 1860.5; DB 2; Length 1011;
 Best Local Similarity 39.6%; Pred. No. 2.1e-101;
 Matches 426; Conservative 168; Mismatches 353; Indels 129; Gaps 27;
 QY 1 MQTSFKKFLS--MILAYSCCLNGGGYAAETMVGQIYDGETLTVSRPVTYVIGDPSGTT 58
 DB 1 MKASLRKFLISTTLTLPYSF-----QAFSLKVVVVPNGTYDG-NLRETPTTITSNEGTT 54
 QY 59 VPSAGELTLNLDNSIAALPLSCFGLNGLSGFTVLGRGHSITPENIRTSNGAALSN---- 114
 DB 55 AILSGNLNLDNSIAALPLSCFGLNGLSGFTVLGRGHSITPENIRTSNGAALSN---- 114
 QY 115 SAADGLFTIEGKELSPNCNLSLAVLPAATNKGSOPTTTTSTPSNGTIYKTDLLLN 174
 DB 115 TPESPPYTIKGVNTLSPNSC---LALMARTTTPATNTTTPV---NPNNGGAFYSKAPVPLEN 168
 QY 175 NKEKFPYSLVSGDGAIDAKSLTVQGTISKLQFQENTAAQDGGACQVTVTSAMANEAP 234
 DB 169 IQNVLTKNNRAADSGGLWVETAGISNKKSMQFLSNVG-ANGGA--INAKSLDVTQCP 225
 QY 235 -TAFVANVAGVGGGIAAVQDQGGVSSSTSTEDPVVSPSRNTAVFEDGNVAVRGGIYS 293
 DB 226 SILFNSNABKGGAIQAVD-----PATNQNTAVRFSNGSVQFDANNAKSGAIYS 279
 QY 294 YGNVAFPLNNGKTLFLNNVSPVYIAAKQPTSGQASNTSNNGYDGGGAIKCKNGAAGSNNSSGVS 342
 DB 280 KGNVDFSNNAQLLIQNNSASP-----EVANTNEVLGGGAIKCKNGAAGSNNSSGVS 328
 QY 343 ----KNGAAGSNNSSGVSFDEGVVFPSSNVAAGKGAIAKGLSVANCGPQVFLRNAND 398
 DB 329 PPPTTNPVFSG-----LTITNQKQILFANNFAATAGGAIYGEKVSITSGKTMFTNNIA 382
 QY 399 NDGGAITLGESGELSLSADYGDIIIPDGNLKTAKENADVNGVTVSSQAIKNGSGGKITT 458
 DB 383 KGGGAIYIPENGELTILSADYGDIIIPDGNLKTAKENADVNGVTVSSQAIKNGSGGKITT 433
 QY 459 LRAKAGHQLFNDPIEMANGNNGNPQAQSSKLLKINDGE-----GYTGDIVF----- 503
 DB 434 LAASGDHKLCFYDPIVITLTPETAPT-NDKTLTINQDKTSSTPTTNYIGTLFSGAVVDSQ 492
 QY 504 -----ANGSTLYQNVTIEQGRIVILREKAKLSVNSLSQ-TGGSLVMEAGSTWDFVTPQPP 557
 DB 493 SASSTTANFESTIYKQVILGGKVLADKASLSVASFTQETDSTILMDNGTTL--LAITEHS 550
 QY 558 QQPPAA-----NQLTLNHLHLSLLANNAVTPNTPPAQDSHPAV 601
 DB 551 HOTPAAGGGGGGGTPTQEBANTDGVISLNLHVNISLSTEQGGALEK----- 599
 QY 602 IGSTTAGSVTISGPIFFEDLDOTAYDRYDMLGSKNOKINVLKQLGTGKPPANAPSDL 658
 DB 600 --KNTDGTITLTHGVSLDDVSGTAYENHDLFNKQVTTINLSTAGDSKTTING-LDLT 656
 QY 659 L-GNEMPKYGYOGSKWLAWDPNTANNGPYTLKATWTKGVNPGPERVASI.VPNSLWGSIL 717
 DB 657 LRGDARFPQYGYOGSKWLAWDPNTANNGPYTLKATWTKGVNPGPERVASI.VPNSLWGSIL 716

```
QY 718 DIRSAHSAIQASVDGRSVCRLVWGSVGNFFYHDDALDGGVRYISGGYISLGANSY-FGS 776
Db 717 DLRSMNALATASCDGFGKGLWVAGISNIFFHDDRSNSVSHGFRRISSGVIIGANSQTWTD 776
QY 777 SMFGLAPTEVFGRSKYDVVCRSNHHACISGVYLSLSTQALCGSYLFGDAFIRASVYFGNQH 836
Db 777 SVFVGAFQCIIPAKSKDYVVSAAKSAQIAGSAYLSVKRQLSNT-IFSSFAARINYSHTNED 835
QY 837 MKTSTVTPAESDVRWNNCLAGEIQAGLPVITPESKLYNELRPVPOAFSVADESPT 896
Db 836 MKRTYTFPEKDGWNNCMLGEIGSLPIVLQITKLHNLQIIPMNVLQGVABHGSFKE 895
QY 897 EGDQARAFKSHLLNLSPVGVKPRCSSTHNPKNYSFMAAYICDAYRTISGTEITLLSHQ 956
Db 896 KLAEARSCSSLRLNLAVPVFKIDRRSHSHGPDYSLAISIPDVWRNRNPGCNTLLLANG 955
QY 957 ETWTTDAPHLARHGVVVRGSMYASLTSNIYVGHGRYERDASRGYGLSAGSRVRF 1012
Db 956 VRMKTPTATNLNRHGLLMQSGTHAVLSNIEIFSHGSCELRSSRNYINVGSKIRP 1011

RESULT 4
ID Q5L6J7 CHLAB PRELIMINARY; PRT; 1024 AA.
AC Q5L6J7
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polymorphic outer membrane protein.
GN Name=pmp7G; OrderedLocusNames=CAB269;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=826/3;
EX PubMed=15937807; DOI=10.1101/gr.3684805;
RA Livingston M., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Ormond D., Mungall K., Clarke K., Feltwell J., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrall B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation."
RL Genome Res. 15:629-640(2005).
DR EMBL; CR48038; CAH63725.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 1024 AA; 108663 MW; 268B61415C8FD434 CRC64;
```

```
Query Match 33.7%; Score 1776; DB 2; Length 1024;
Best Local Similarity 38.3%; Pred.No.2.2e-96;
Matches 412; Conservative 167; Mismatches 381; Indels 116; Gaps 27;

QY 1 MQTSFHKFPLSMILAYSCSLNGGGAABIMVPOGIYDGETLTVSPFYTVIGDPSGTVF 60
Db 1 MKASFRKELVSTTLTPC---SFQAFSLIIVPNTYNGD-LREMFPTIISNEGITAI 56
QY 61 SAGBLTKNLNLSAALPLSCFNLGSLFTVLGRHSITPFRITSTNGAALS-----NSA 116
Db 57 LSGNLNLNLSNMAATPSSCFPNSAGSMWIIKGHDVFTPLRLTSVNGAALSILITSSP 116
QY 117 ADGLFTIEGPKELSPBNCNSLILAVLPAATTNGSGQTPTTTPSPNGTIYSKTDLLNNE 176
Db 117 E8F5YITIGVTGTFSCNSCALL-----GRNNSNLSILSPKGAAYYSKSPFFKNIQ 166
QY 177 KFSFVSNLVSGDGAIDAKSLTVQGISKLCVFOENTAOADGACQVTSFSMANEAP-I 235
Db 167 NLIIFKNCADANGGALMQQVVDISNVTSLFLSNVG-ANGAIGASTSLN--VTRCP8I 223
QY 236 AFVANAVRGGGIAAVDQGGQGVSSSTSTEDP-----VVSFSRNTAVFEGNVAR 286
Db 224 LFRNSASRLRGAI-----HSVNPQTPPPPPPPGNGVINTVNFSDNGSVQVDFSNNAK 276
```

```
QY 287 VGGGIYSYGNVAFPLNNGKTLFLNNVASPVYIAAQPTSGQASNTSNNYGDGAIFPKNGA 346
Db 277 SGAIIYSKGNINFNQVOLLQNTASP-----EIGDSNEVLGGGAIYC---T 322
QY 347 QAGSNNSGSVDFDEGV-----VFFSSNVAAGKOGAIIYAKKLSVANCPVPVFLRNANDG 401
Db 323 QATATKKAAPVFTGLTITNQHDIFPANNFAANAGGAIYGEKVSITSSGRTVFTNTAKTG 382
QY 402 GAIYLGSGELSLADYGDIIIPDNLKRTAKENADVNGVTVSSQAISMGGGKITTLRA 461
Db 383 GAIYIADNGELSLADYGNMTFYDNL-----NTNGSGSPKRNNAVTLGKGATIKLAA 433
QY 462 KAGHQLLPNDPIEMANGNNOPAQSSKLLKINDGSG-----VTGDIVP-----AN 505
Db 434 SGDHLKCYFDIVTTLPEKAP-NGSNTLTINPDGNSAPFTNYIGTVLFSGAHANAESAP 492
QY 506 GSSTLYQNVITTEQRIIVLREKAKLSVNSLSQTGSLYMEAGSTWDFVTPQ----PPOOPP 561
Db 493 NASTIYQKVLGGKVLADKASLSVVSFQDSISILLMDNGTSLTITEHSYATSSA 552
QY 562 AANQLITLSNHLSSLANNA-VTNPPPTNPP-----AODSHPAVIGS-T 605
Db 553 AAASVVTAAAEVSSVTSSATAAASVNTQANSDDGVVSIKDLHNLGSLTQDGEKAKETKN 612
QY 606 TAGSVTISGRIPFPEDLDDTAYDRYDMLGSNOKINVLKQLGKTGPAN-APSDTL---GN 661
Db 613 TSGTITISEHISLDDVSGNAYENHDIFNRN-TVTLKVLSTAGDNKITESDLQLTFRGD 671
QY 662 EMPKYGYQGSWKLAWDPNTANN---PYTLKATWTKTYGPNPGRVVASLVPNSLWSIL 717
Db 672 ADPOGYQGSWKLAWENGT--NGDAKKVKLAKATWTKTFVPSPERQASLVPNSLWGAFI 729
QY 718 DIRSAHSAIQASVDGRSVCRLVWGSVGNFFYHDDALDGGVRYISGGYISLGANSY-FGS 776
Db 730 DLRSMNALATASCDGFGYKGLWVAGISNFFHDDRSNSVSHGFRRISSGVIIGANSQTWTD 789
QY 777 SMRGLAPTEVFGRSKYDVVCRSNHHACISGVYLSLSTQALCGSYLFGDAFIRASVYFGNQH 836
Db 790 SVFVGAFQVFGSKDYVVSAAKSAQIAGSAYLSIKRPLSNT-IFTTFAARINYSHTNED 848
QY 837 MKTSTVTPAESDVRWNNCLAGEIQAGLPVITPESKLYNELRPVPOAFSVADESPT 896
Db 849 MKRTYTFPEASGNWNNCMLGEVGSPLIVLQITKLHNLQVFPVFNVLQGVABHGSFKE 908
QY 897 EGDQARAFKSHLLNLSPVGVKPRCSSTHNPKNYSFMAAYICDAYRTISGTEITLLSHQ 956
Db 909 QLAEARSCSSLRLNLAIPOGFKIDRRSHSVDFYSLAVSYVPDVWRNRNPGCNTLLLANG 968
QY 957 ETWTTDAPHLARHGVVVRGSMYASLTSNIYVGHGRYERDASRGYGLSAGSRVRF 1012
Db 969 ARMKTPTATNLDRQLLIGSSHTAVNNNIEIFSHGSCELRSSRNYINVGSKIRP 1024

RESULT 5
ID PMP13 CHLPN STANDARD; PRT; 973 AA.
AC Q92896; Q9K2A1; Q92410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmp13 precursor (Polymorphic membrane protein 13) (Outer membrane protein 14).
DE Name=pmp13; Synonyms=omp14;
GN OrderedLocusNames=CP0453; CP0299; CP80470;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
```

"Comparative genomes of Chlamydia pneumoniae and C. trachomatis."; Nat. Genet. 21:385-389(1999).

[2]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

STRAIN=AR39;

MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397; Read T.D., Bruham R.C., Shen C., Gill S.R., Heidelberg J.F., White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J., Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C., Dodson R.J., Gwin M.L., Nelson W.C., DeBooy R.T., Kolonay J.F., McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.; "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39."; Nucleic Acids Res. 28:1397-1406(2000).

[3]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

STRAIN=J138;

MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311; Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ihil K., Hattori M., Kuhara S., Nakazawa T.; "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA."; Nucleic Acids Res. 28:2311-2314(2000).

[4]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

STRAIN=TM-13;

Gen M.W., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P., Schneider S., Pohl T., Essig A., Marre R., Melchers K.; "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis."; Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

[5]

NUCLEOTIDE SEQUENCE OF 1-262.

STRAIN=CWL029 / VR1310;

MEDLINE=20007584; PubMed=10539856; Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P., Madsen A.S., Knudsen K., Falk E., Birkelund S.; "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenesis."; Am. Heart J. 138:S491-S495(1999).

CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (potential).

CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.

CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AE001629; AAD18595.1; -; Genomic DNA.

EMBL; AE002191; AAF38156.1; ALT_INIT; Genomic DNA.

EMBL; BA000008; BAA39860.1; -; Genomic DNA.

EMBL; AE017158; AAP98401.1; -; Genomic DNA.

EMBL; AJ133034; CAB37074.1; -; Genomic DNA.

PIR; B86547; B86547.

PIR; F72076; F72076.

PHCI-2DPAGE; Q92896; -.

TIGR; CP0299; -.

InterPro; IPR005546; Auto transpbteta.

InterPro; IPR011427; ChlamPMP.M.

InterPro; IPR003368; Chlamydia_PMP.

Pfam; PF03797; Autotransporter; 1.

Pfam; PF02415; Chlam PMP; 8.

Pfam; PF07548; ChlamPMP M; 1.

TIGRFAMs; TIGR01376; POMP repeat; 6.

KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.

SIGNAL 1 24 Potential.

CHAIN 25 973 Probable outer membrane protein pmp13.

CONFLICT 258 258 N -> Y (in Ref. 5).

SQ SEQUENCE 973 AA; 102762 MW; E02AG69FG11DEBFF52 CRC64;

Query Match	25.9%;	Score 1365.5;	DB 1;	Length 973;
Best Local Similarity	33.5%;	Pred. No. 4.2e-42;	Indels 103;	Gaps 27;
Matches 350;	Conservative 171;	Mismatches 720;		
Qy	1	MOTSHPKFLSMILAYSCSLNGGGYAAIMVPGQIYDGETLTVSPFYTVIGDPGSGTTFV	60	
Db	1	MKTSIRKFLISTLTACPAST---AFTVEVIMPESENFDGSSGKI-FPYTILSDPRGTLCI	56	
Qy	61	SAGBLTKNLNDNSIAALPLSCFQNLGSGTVVLGRGHSILFENIRTSITNGAALSNSAADG-	119	
Db	57	PSGDLVIANDLNAISRTSSSCFSNRAGALQILKGGVFSFLNTRSSADGAASISVITQNP	116	
Qy	120	---LFTIEGPKELSPSCNLSLLAVLPAATNKGSGQTPTTSTPSNGTIVYSKTDLLILLNNE	176	
Db	117	ELCPLTSPFGSPSOMIFONCESLT-----SDTSASNVI PHASAIYATTMPLFTNND	165	
Qy	177	KPFYFYNLWSDGCGAIDAKSLTVQGISKLVCFOENTAQADGCAQCVVTSFSAMANEAPIA	236	
Db	166	SIILPOYNRSAGFAALRGTSITIENTKSLFPNGNGSISNGGALTGSAATLNINNSAPVI	225	
Qy	237	FVANVAGVRGGGIAAVQDGOOGQVSSSTFEDPVVFSRNTAVAPDGNVAVRGGIYISGN	296	
Db	226	FSTNATGIYGGAIVLT-----CGSMLTS-----GNLSGVLFPVNSSRSGGAIYANGN	272	
Qy	297	VAFLNNGKTLFANNVASPV-YIAAKQPTSCQASNTSNNGYDGGAI FCNKGCAQAGSNNSGS	355	
Db	273	VTFSSNLDLTFQNTTASQPNSLPATPPPTTAVTP-LLGYGGAIFCTPPATPPPTGV-S	330	
Qy	356	VSPDGEVGFSSNVAAGKGAIVAKCLSVANCGPVQFLRNIAINDGGAIVLGSGLSLS	415	
Db	331	LTISGENSVTFLENIASEQGALYKKLISDSNKSITIFLGNWAGKGAIALPESGELSLS	390	
Qy	416	ADYGDIIIFDGNLKRITAKENAADVNGVTVSQAISMGSGGKITTLRAKAGHQILFNDPIEM	475	
Db	391	ANQGDILFNKLSITSG-----TPTRNSIHFGKDAKFATLGATGQVTLFYDPDI--	439	
Qy	476	ANGNNQPAQSCKLLKIN-----DGEYTGDIVP-----ANGSTLYQNVTIE	517	
Db	440	TSDDLASAASAAATVVVNPFKASDG-AYSGTIVSGETLTATREATPATNATSTLNQKLELE	498	
Qy	518	QGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQPPQPPQPPAANQILITSLNLHLSL	576	
Db	499	GGTIALRNGATLVNHTTQDEKSVVIMDAGTT--LATNGANNTDGA---ITLKNKLVNL	553	
Qy	577	SSLIANNNAVTPNPTNPPAQBDSHPAIG-STTAGSVTISGPIFFBDLDDTAYDRYDWLGSN	635	
Db	554	DSLDTKA-----AVNVQSTNGALTISGTLGLVKNSQDCCDNHGMFNKO	598	
Qy	636	-QKTNVLKQLGKTPKANASDIL-TLGNEMPKYQYQGSWKLANDPNTANNPGPYTLKATWT	693	
Db	599	LQQVPILLELKATSNVTVTTPDPSLGTNGYQOSPYGYQGTWETIDTITTT-----HVTYGNWK	653	
Qy	694	KTYNPNPGRVASLVPNSLWGSILDIISAHSALQASVDGRSY-CRGLWVGVSVPFFYHDR	752	
Db	654	KTYGLPHERLAPILPNSLMANVILDKAVSOA--SAADGEDVPGKQSLTIGITIFFHANH	711	
Qy	753	DALGQGVRYISGQYSLGANSY---FGSSMFLGAFTEVFGRSKQYVVCNRSHHACIGSYVL	809	
Db	712	TGDARSYHMGGGYLI--NTYTRITPDAAALSLGFLQFTKSKOYLVGHSNVYFATVYS	769	
Qy	810	STQOALCG-SYLFQDAPFIRASYFGFNQHKMTSYTFABESDVRWNONCLAGIAGLPIVI	868	
Db	770	NITKSLFGSSRFPFGGTSRVTVYSRNEKVKTSYTKL PKGRCSWNNCWLGLLEGKLPITL	829	
Qy	869	TPSKLYLNLRLPFVQAEPSYADHESPTTEGQOARAFKSGHLLNLSPVGVKVFDRCSSTHP	928	
Db	830	SSRIINLQKIIIPFKAEYAYATHGCIQENTPEGRIFGHGHLINLNAVPGVGFGRKNSHRP	889	
Qy	929	NKYSFMAAYICDARTISGTTETLISHQETWTTDAFLHARGVWVRGSGMYASLTSNIENV	988	
Db	890	DFYTIIVAYADVYRHPDCTTLPINGATWTSIGNLSTRLLYQOASHTSVNDVLBIP	949	
Qy	999	GHGRYEYRDSARGYGLSAGSRVF	1012	

SQ SEQUENCE 926 AA; 98439 MW; 3E75B52P594750F CRC64;
Query Match 21.3%; Score 1122; DB 2; Length 926;
Best Local Similarity 29.7%; Pred. No. 1.1e-57;
Matches 323; Conservative 152; Mismatches 376; Indels 236; Gaps 34;
QY 1 MTSFHKPFLMILAYSCSNGGVAABIMVPOGIYDGETILTVSPFYTVI-----51
DB 1 MRPSLYKILISSTL-----TLPISEHFSQALAEVALTQE 34
QY 52 -----GDPSTGTVFSAGBLTKNLDSI-----AALPLSCFNGLLGSGTVLGRHS 97
DB 35 SILDANGAFSPQSTAGG-TIYNVESDISVDVGTAAALASSAFVQTADNLTFKGNHS 93
QY 98 LTFENIRTSNGAALSNAADGLFTIEGFKELSPSNCNLLAVLPAATNKGSTPTTTS 157
DB 94 LSIITNANAGANPAGINVTADKILTLTDFSKLSFKEC-----PSSLVNTG-----138
QY 158 TPSNGTIYKTDLLLNNEKPSFYNLVSGDGAIDAKSLITVQGISKLCVFQENTAAQDG 217
DB 139 ----KGAMKGGGALNANNASILFDQNYSAENGGAISCKAFSLTSGSKEISFTNSTAKG 195
QY 218 GACQVTVFSAMANEAPIAFVANVAGVGGGIAAVDQGGQVSSSTSTEDPVVFSRNTA 277
DB 196 GA-----IAATGIAHLSDNQ-----GTIRFSGNTA 220
QY 278 VEPDGNVAVGGIYSGNVAFNLNGKTLFLNNVAFVIAAKQPTSGQASNTSNNGYDG 337
DB 221 V-----NSGGAVYSEASMTIAGNNHVAFSNNAVS-----GSSDGGC 256
QY 338 GAIFCKNGAQSNGSGVSFDGEGVFFSSNVAAGKGAIAKLSVANGCGVQPLRN- 396
DB 257 GAHCSK-----TGSAPTLIRDNKVLIFEENTSSAKGAIYDILKILISGGTAFINK 311
QY 397 ---IANDGGAIFYGESGELSADYDIIIPDGNLKR TAKENAADVNGVTSSQAIMSGG 453
DB 312 VTHATPKGAIGTAANGECSLTAHEGDIITPDNNLMAT-QDNA-----TIKRNAINEGN 364
QY 454 GKITTIRAKAGHOILFNDPIEMANGNQPQOSKLLKINDGEG---YTGDIVFA-----504
DB 365 GKPVNLRASAGKTIISFYDPIIV-EGN-----AADLLTLNKAEGDKTYNGRIIFSGEKLTE 418
QY 505 -----NGSSTLYQNTVIEQGRVILREKAKLSVNSLSQNGSI-VMEAGSTWDFVTPQ 556
DB 419 EQAAVADNLKATFTQPTTLAAGELVLRSGVEVEAKTVVQTAGSLILMDAGT-----469
QY 557 PQPPAANQLITLNLHLSLSSILANNAVTPNPPAQDSHP-AVIGSTTAG-SVTISG 614
DB 470 -----KLSAKTEDATLNLAINPNTLDGKFAVVDVAAGKNTLSG 511
QY 615 PIFPEDLDDTAYDRYDWLGSNQKIN-----VLKQLQ-----GTPKANAPSDLTGLNEMPKYG 667
DB 512 AIGVIDPTGKPYE-----NHKLNLDLALGGIQLSGKSVTTTNNVPSHV-VGVAETHYG 563
QY 668 YQGSWKLAWDNTANGPYTLKA---TWTKYGNPGRVAVSLVNSLWGLSIDIRSAHSA 725
DB 564 YQGNWSVSKDN-NSDPKTKQTALFTWNKTYGVNPPERRAPLVNLSWGSFIDURSIOV 622
QY 726 IQASVDG-RSVCRLWVGVSNNPFDHRLDAGQGRYVTSIGSYLGANSYFPG-SSMFGIAP 783
DB 623 LERSVDSILETRGLWVGSGNPFHKORNAENRFRHISGTVLGATNTSREDSLSVAF 682
QY 784 TEVFRSKDYVVCRSNHACITGSYV-----LSTQQALCGSY-----LFGDAF 825
DB 683 QQLFAKOKDYLSKNAANVAGSVYQHVSKFDDLTFLFNGPNTCCSGFSKEIPIFLDAQ 742
QY 826 IRASVFGNQHKMSTSYTPAESDVRWNCLAGIAGLPIVITPSKLYLNLRLPFPQAB 885
DB 743 I--TYCHTANNMTTSYTDYPEVKGSGWGNDTLGLTLSTSVPIPVFSSSIF-DSYAPFAKLQ 799
QY 886 PSYADHESFTEGQOARAFKSHLNLNLSPVGVKFDRCSTHPNKYSMAAYICDAYRTI 945
DB 800 VVYAHQDDFKBPTTEGRVFPSSDLLNVSVPGIKPFKLSYGERSAYDLTLMYIPDVTRHN 859

QY 946 SGTETILLSHOETWTTFDAPHLARHGTVVGRSMYASLTSNIEVYHGRYEVYDASRGYGLS 1005
DB 860 PSCMTGLAINDSVSLTATNLARQAFIVRAGNHIALTSGVEMFQFGFELRSSSRNVD 919
QY 1006 AGSRVP 1012
DB 920 LGAKVAF 926

RESULT 8

Q5L6J3 CHLAB PRELIMINARY; PRT; 926 AA.
AC Q5L6J3;
DT 01-FEB-2005 (TREMELrel. 29, Created)
DT 01-FEB-2005 (TREMELrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
DE Polymorphic outer membrane protein.
OS Name=pmp14G; Synonyms=pomp98A; OrderedLocusNames=CAB282;
GN Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S26/3;
RX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640(2005).
DR EMBL; CB848038; CAH63732.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 926 AA; 98440 MW; F81C2A79A3D575D1 CRC64;

Query Match 21.3%; Score 1122; DB 2; Length 926;

Best Local Similarity 30.3%; Pred. No. 1.1e-57;

Matches 325; Conservative 151; Mismatches 389; Indels 208; Gaps 35;

QY 1 MTSFHKPFLMILAYSCSNGGVAABIMVPO-GIYDGETILTVSPFYTVI-GDPSTGTV 59
DB 1 MRPSLYKILISSTLTP-ISPFHSQLHAEVALTQESILDAN-----GAFSPQST 48
QY 60 PSAGELTLKXLDNSI-----AALPLSCFNGLLGSETVLGRHSLTFENIRTSNGAA 111
DB 49 STAGG-TIYNVESDISVDVGTAAALASSAFVQTADNLTFKGNHLSITNANAGANPAG 107
QY 112 LNSAADGLPTTEGFKELSPSNCNLSILLAVLPAATNKGSTPTTTPSPNGTIYKTDLL 171
DB 108 INVNTADKILTLTDFSKLSFKEC-----PSSLVNTG-----KGAMKSGGALN 149
QY 172 LLANKEFSPYNLVSGDGAIDAKSLITVQGISKLCVFQENTAAQDGAQVTVTSFSMAN 231
DB 150 LANNASILFDQNYSAENGGAISCKAFSLTSGSKEISFTTNTAKGGA-----197
QY 232 EAPIAFVANVAGVGGGIAAVDQGGQVSSSTSTEDPVVFSRNTAVEPDGNVAVGGGI 291
DB 198 -----IAATGIAHLSDNQ-----GTIRFSGNTAV-----NSGGAV 227
QY 292 YSYGNVAFNLNGKTLFLNNVAFVIAAKQPTSGQASNTSNNGYDGAIFCKNGAQAQGSN 351
DB 228 YSRASMTIAGNNHVAFSNNAVS-----GSSDGGCAGAHCSK-----TG 265
QY 352 NSGSVSPDGEVGVFFSSNVAAGKGAIAKLSVANGCPVQFLRN-----IANDGGAIFYL 407
DB 266 SAPTLIRDNKVLIFEENTSSAKGAIYDILKILTSGGTAFINNKVTHATPKGAIGIA 325
QY 408 ESELSLSADYDIIIPDGNLKR TAKENAADVNGVTSSQAIMSGGKLTTLRAKAGHOI 467
DB 326 ANGECSLTAHEGDIITPDNNLMAT-QDNA-----TIKRNAINEGKGFVNLRAASGKI 378


```
QY 468 LFPNDPIEMANGNQPAQSKLLKINDGBG---YTGDIIVFA-----NGSSTLY 511
DB 379 SFYDPIV-EGV-----AADLLTKAAGDKTYNGRIIFSEKLTTEEQAADVADNKTFTT 432
QY 512 QNVITIEQRIVLREKAKLSVNSLSOTGSL-YMEAGSTWDFVTPPOPPQPPAANQLITLS 570
DB 433 QPITLAAGELVLSRGSVEAKTVQTAGSLILMDAGT----- 469
QY 571 NLHLSLSLLANNVNTPTTPPADSH-PAVIGSTTAG-SVTISGPIPFDDLDYADR 628
DB 470 ----KLSATEDATLNLAINPTLIDGKFAVDAVAAGKNTLSGALGVIDPTGKPYE- 524
QY 629 YDMLGSNOKIN----VLKQL---CTKPPANAPSDLTGNMPPKYGYQGSWKLAWDPNTA 681
DB 525 -----NHKLNDTLALGGIQLSGKGSVTTTNPVSHV-VGVAEETHYGYGNGNSVSWKDN- 576
QY 682 NNGPYTLKA--TWITGYNPGPVERVASLVPNSLMGSIIDIRSAHSAIQASVDG-RSYCRG 738
DB 577 NSDPKTQTATFTWNTKGYVNPERRAPLVLNSLWGFIDLSIQDVLERSVDLSILETRG 636
QY 739 LWVGVSVPFFYHRRDALGGVRYISGGVSLGANSYFG-SSMFGLAFTVFGSKDYVVCVR 797
DB 637 LWVGIGNFFPHKDRNAENRKHRISSGYVLGATTWTSREDSLSVAFQCFANDKDYLVSK 696
QY 798 SNHACIGSVY-----LSTQALCGSY-----LPGDAFIRASYGFGNQHMKT 839
DB 697 NAANYAGSVYVYQHVSKFDDTLRLPFGNPTCCSGFSKSIPIFLDAQI--TYCHTANNMTT 754
QY 840 SYTFAEESDVRWNNCLAGEIGAGLPIVITPSKLYNELRPFVQAEFVSADHESPTERGD 899
DB 755 SYTDYPEVKGSGNDTLGLTSTVPIVPSSIF-DSYAPFAKLQVYAHQDDPKPTT 813
QY 900 QARAFSGHLLMLSPVGVKPDRCSTHPNKYSFMAAIVCAIRYTSITETTLSSHQWTW 959
DB 814 EGRVPESDLLNVSPVIGIKPEKLSYGERSAYDLTMTYIPDYRHNPCMTGLAINDVSW 873
QY 960 TTDAFLARHGTVRSGVYSATISNIEVYHGRVEYRDASRQYGLSAGSRVRF 1012
DB 874 LTTATNLARQAFIVRAGNHIALTSGVEMPFGFPELRSSRNYNVDLGAQVAP 926

RESULT 9
PMP10_CHLPPN
ID FMP10_CHLPPN STANDARD; PRT; 928 AA.
AC Q9RB65; O86163; Q9RB64; Q986P2;
DT 16-OCT-2001 (Rel. 40, Created)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable outer membrane protein pmp10 precursor (Polymorphic membrane protein 10) (Outer membrane protein 5).
GN Name=pmp10; Synonyms=omp5; OrderedLocustNames=CP0303, CpB0467;
OS Chlamydia pneumoniae (Chlamydia pneumoniae pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83358;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=105339856;
RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity."
RL Am. Heart J. 138:8491-8495(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
```

```
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE OF 1-914.
RC STRAIN=CWL029/VR-1310;
RX MEDLINE=99081766; PubMed=9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer
RT membrane proteins of Chlamydia pneumoniae."
RL Infect. Immun. 67:375-383(1999).
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AJ133034; CAB37071.1; -; Genomic_DNA.
CC EMBL; AB002192; AAF38160.1; -; Genomic_DNA.
CC EMBL; BA000008; BAA98657.1; -; Genomic_DNA.
CC EMBL; AB017158; AAP98398.1; -; Genomic_DNA.
CC EMBL; AJ001311; CAA04671.1; -; Genomic_DNA.
CC PIR; G81591; G81591.
CC PIR; G86546; G86546.
CC PHCI-2DPAGE; O86163; -.
CC TIGR; CP0303; -.
CC InterPro; IPR005546; Auto_transp_beta.
CC InterPro; IPR006315; Auto_transporter.
CC InterPro; IPR011427; ChlamPMP_M.
CC InterPro; IPR003368; Chlamydia_PMP.
CC InterPro; IPR006626; Pdh1.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; Chlam_PMP; 7.
CC Pfam; PF07548; ChlamPMP_M; 1.
CC SMART; SM00710; Pdh1; 4.
CC TIGRPFAMS; TIGR01414; autotrans_bar1; 1.
CC TIGRPFAMS; TIGR01376; POMP_repeat; 5.
CC Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 928 Probable outer membrane protein pmp10.
FT CONFLICT 305 305 Missing (in Ref. 4)
FT SEQUENCE 928 AA; 97230 MW; 0590D5206A1DD0E1 CRC64;
SQ
Query Match 20.9%; Score 1101; DB 1; Length 928;
Best Local Similarity 30.5%; Pred. No. 1.9e-56;
Matches 324; Conservative 150; Mismatches 402; Indels 188; Gaps 34;
QY 1 MQTSPHFPPFLSMILA-YSCSLNGGGAALIMVQGIYDGETLTVSFYVIGDPSGTTV 59
DB 1 MKSQFWLVLSTLACFTSCSTVFPA-PAENIGPSDFSGSTNVTGT--YTPKNTTGTIDY 57
QY 60 FSAGELTKNLNDNSIAALPLSCFNLGSLFTVLRGHSHTFENIRTSNTGALSNSAAG 119
```

Db	58	TLTQDITLQNLGDS--AALTGKCFSDTTESLSPAGKGYSLSPFLNKSSAEGAALS-VTTDK	115
Qy	120	LETFEGFKELSPNCSNLLAVLPAAATTKGSGTPTTTSPNGTIYSKTDILLLNKEFS	179
Db	116	NLSLTCFSSLT-----LAAPSSVITTPSG-----KGAVKCGDGLTFDNNGTIL	159
Qy	180	FYSNLVSGDGGDAIDAKSLTVQIGSKLCVFOENTAOADGACQVVTFSFMANEAPIAFA	239
Db	160	FQDYCEENGGAIISTNLSKNSGISIFEGNKSSA-----	195
Qy	240	NVAGVRGGIAAVQDQCGVSSSTEDPVVFSRNTA-VEPDGNAR-VGGGIYSYGNV	297
Db	196	--TGKKGGAICA-----TGT-----VDITNNTAPTLSNIIAARAAGAINSTGNC	238
Qy	298	APLNGKTLFLNNVASPVVIAAKQPTSGQASNTSNNGYGGGAIFCKNGAQAQSNNSGSVS	357
Db	239	TITGNTSLVFSN-----SVTATAGNGGAL-----SGDADVT	270
Qy	358	PDGEGVVPFSSNVAAGKGAIIYAKKLIVAN--CGPVQPLRNI-----ANDGGAIVLBSG	410
Db	271	ISGNOSVTPSGNQAVANGGAIYAKKLITLASGGGGISFNNIVQGTAGGGAISILAAG	330
Qy	411	ELSLADYGDIIIPDGN-----LKRITAKENAADVNGVTSSQAISMGSQGIITTLAKAGH	465
Db	331	ECLSAERAGDIIFNGNAIVATPQTTRNSIDI-----GSTAKITNLRAISCH	378
Qy	466	QILFNDPIEMANGNNOPOASSKLLKINDGEG-----YTGDIVPA-----NGS	507
Db	379	SIFPYDPITA-----NTAADSTDTLNLKADAGNSTDYSGISVFSGEKLSDEAKVADNLT	434
Qy	508	STLYQNTVTEQGIIVLREKAKLSVNSLSOT--GGLYMEAGSTWDVFTVPPQPPAANQL	566
Db	435	STLKQPVITTAGNLVLRGVTLDTKGTQTAGSIVMDAGTTL-----KASTEE	483
Qy	567	ITLSNLHLSSLANNNAVTPNPPADSHPAVIGSTTAGSVTISGPIFFEDLDDTAY	626
Db	484	VILTGLSIPVDSL-----GEGKKVIVIAASAASKVALSGPILLLLDNQNAV	529
Qy	627	DRYDWTGNSQKINVLKLO--LGTKPPANAPSDLTLGNEMPKYGVQGSWKLAWDNTANNP	685
Db	530	ENHD--LQKTDQDFSVQLSALGTATTDDPAVPTVATP--THYGYQGTWGMTVDDDTAST--P	596
Qy	686	YTLKAT--WTKTYNPGPERRVASLVNSLWGSILDIRSAHSAIQASVDGRSVCRLWVSG	743
Db	587	KTKTATLAWNTGYLPNRPQGPLVPSNLWGSFSDIQAIQGVIERGALTLCSDRQFWAAG	646
Qy	744	VSNFFYHDDRDLAQGYRISGGYSL--CANSYFGSSMFGLAFTBVFGRSKDYVVCRSNHRA	802
Db	647	VANFLDKKGEKRYKTHKSGGYAIGGAQTCSENILISPACQLFGSKDKDPLVAKNHTDT	706
Qy	803	CIGSVYLS-----TQOALCGSYLFGDAFI--RASYPGNGHMKTSYTFBESDV	849
Db	707	YAGAFYIHITECSGFGICLLDLKPGSWSHKPLVLEGQLAYSHVSNDLTKYTAYPEVK	766
Qy	850	RWNCLAGBIGAGLPIVITPSKLY--LNELRPFQABFSDYADHESFTTEGDQARAYKSGH	908
Db	767	SWGNNAPNMLGASSHSY--PEYLHCPTDYAPYIKMLTYIQDQSFSEKTEGREGSDDSN	824
Qy	909	LNLISVPVGVKFDRCSTHFNKYSFMAAYICDAVRTISGTEITLLSHOETWTITDAPHAR	968
Db	825	LFNLSLPIGVKFKFSPDCNDFSDLTLSYVPDLIRNDPKCTTALVTSIGASWETVANNLAR	884
Qy	969	HGVVVRGSMYASLTSNIEVYGHGRYERDASRGYLSAGSRVRF	1012
Db	885	QALQVRAGSHYAPSPMEFVLQGFVEFVRGSSRIYVNDLGGKQF	928
RESULT 10			
ID	Q823W9	CHLCV PRELIMINARY;	PRT; 926 AA.
AC	Q823W9		
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	


```

DR EMBL; AB001627; AAD18590.1; -; Genomic DNA.
DR EMBL; AE002193; AAP38164.1; -; Genomic DNA.
DR EMBL; BA000008; BAA38654.1; -; Genomic DNA.
DR EMBL; AE017158; AAP98394.1; -; Genomic DNA.
DR PIR; A81591; A81591.
DR PIR; D72078; D72078.
DR PHCI-2DPAGE; Q92393; -.
DR TIGR; CP0307; -.
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 930 Probable outer membrane protein pmp8.
FT VARIANT 177 177 T -> A (in strain AR39, strain J138 and strain TW-183).
SQ SEQUENCE 930 AA; 97670 MW; 46A9B5E3BB913C4C CRC64;

Query Match 20.5%; Score 1079; DB 1; Length 930;
Best Local Similarity 28.3%; Pred. No. 3.8e-55;
Matches 303; Conservative 160; Mismatches 407; Indels 202; Gaps 29;

QY 1 MQTSPHKFPLSMILAYSCSLGGGYAAEI-MVPOGIYDGETLTVPSPYTVIGDPGGTTV 59
DB 1 MKPIPLKLLISLTIV-TPIILSIATYGDASLSPTDSFDGAGGSTFTPKST-ADANGTNY 58

QY 60 PSAGELTKLNDNSIAALPLSCFNGLLGSFTVLGRGHSLLTINERTSNGAALNSAADG 119
DB 59 VLSGNVYINDAGKG-TALTGCGCTETTGDLTFTGKYSFSFNTVDAGSNAGAAASTTADK 117

QY 120 LPTIEGRKELSPNCNSLLAVLPAATTNKGQTPTTT---STPSNGTIYSKTDLLLLNNE 176
DB 118 ALTFPGPNLSF-----IAAPGTVASGKSTLSSAGALNTDGTTLFQNV----- 164

QY 177 KFSFYNLSVSGDGAIDAKSLTVOQISKLCVFOENTAOADGACOVVTSFSAMANEAPIA 236
DB 165 -----SNEANNNGGAILTKTSLISGNTSSITFTNSAKLGGALYSSAAASISGNTGOLV 219

QY 237 FVANVAGVRGGIAAVQDQGGVSSSTEDPVVSPSNTAVFPGNVARVGGGIYSYGN 296
DB 220 FMNN-KGETGGALGFE-----ASSSITONSSLPFGSGNTATD----- 255

QY 297 VAFLNNGKTLFLNNVASPVYIAAKOPTSGQASNTSNVYDGGAIKCKGAAGSNNSSGV 356
DB 256 -----AA-----GKGAIYCEK-----TGETPTL 274

QY 357 SPDGBGVVFFSNVAAKGGAIAKLVANCGVPQFLRN-----IANDGAIYLGESGE 411
DB 275 TISGNKSLTPFENSSTVGGGAIKCHGLDLSAAGPTLFNNRCNGVTAAGKGAIAIADSGS 334

QY 412 LSLGADYGDIIFDGNLKRKAKENAADVNGVTVSSQAIMSGSGKITTILRAKAGHQILFND 471
DB 335 LSLSANQGDITFLGN-----TLTSTAPTSTRNAIYVLGSSAKITLRLAAQGSQSIYFD 387

QY 472 PIEMANGN-----NQAQSKLLKINDGEGYTDIVFA-----NGSSTLY 511
DB 388 PIASNTTGASDVLITINQPDNSPL-----DYGSTIYFSGEKLSDRAKAADNFTSLK 440

QY 512 QNVITEQRIVLIREKAKLSVNSLSGTGGSLYMEAGSTWDFVTPPOPPQPPAANOLITLSN 571
DB 441 QPLALASTLALKGNVELDVNGFTGTGSTLL-----MQPGTKLKADTAISLTK 490

QY 572 LHLSSLSSLLANNAVTPNPPAODSHPAVIGSTTAG---SVTISGPIFFEDLDDTAYDR 628
DB 491 LVVDLSALEGNKSVS-----IETAGANKTITLSPLVFQDSSGNFYE- 532

QY 629 YDMLGSNQKINVLKQLGTGP-----PANAPSDLTIG-----NEMPKYGYQGSNKL 675

```


Db	389	ES	-----QEDPATNAPATOGGAKEEAANSFDIILKQNVNITFSSNSKTAG	436	ID	MP7_CHLPN	STANDARD;	PRT;	936	AA.
Qy	247	GGIAA	VQDQGVSSSTEDPVVSFNTAVEFGNVARVGGGIYSYGNVAFNLNGKTL	306	DT	Q92898	Q9JSA42; Q925Q4;			
Db	437	GAIRV	---SGSATIENTGT---CTFSNNNAKE-----OGGALVNGNCIDITGNKVV	483	DT	16-OCT-2001	(Rel. 40, Created)			
Qy	307	FLNNVASPV	-----YIAAKQPTSGQASNTSNNGYDGGAIFC-----	342	DE	13-SEP-2005	(Rel. 48, Last sequence update)			
Db	484	PSGNOAQEI	PAPSIYVVEAAVEEPVEAK-----GSGGAIYCVKAPISIPATPIQPP	535	DE	Probable outer membrane protein pmp7 precursor (Polymorphic membrane protein 7) (Outer membrane protein 12).				
Qy	343	-----XNGAQA	GSNNSS-----GSVFDGEGVVPFSSNVAAGKGAIIYAKKLS	384	GN	Name=pmp7; Synonyms=omp12; OrderedLocusNames=CpN0445, CpB0462;				
Db	536	KQTLPSL	SAALSGETHAKVAQKEBSPDPCLTISGNASVIPDNNSSTVTGGAIHAKKV	595	OS	Chlamydia pneumoniae (Chlamydophila pneumoniae).				
Qy	395	VANCPVQ	FLRNANDGGAIYLBESGELSLSDYDGIILFDGNLKEKTAKEN-----	434	OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.				
Db	596	LSSSGN	MTFSSNKGGAIIYADGGDISITATTGSIIPQGN-KVTAADSIPLTKYKIEA	654	OX	NCBI_TaxID=83558;				
Qy	435	---AADVNG	VTVSSQA-----ISMGSGGK	455	RN	[1]	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
Db	655	IAAESIE	BEKTFPSQASGSASAGTSAPFTLANKAETPAESQAKENSKPTCSNHLGSGAK	714	RP	STRAIN=CWL029;				
Qy	456	ITTLRAK	AGHOILFNDPI-----EMANGNN-----QPAQS	485	RC	MEDLINE=92026606; PubMed=10192388; DOI=10.1038/7716;				
Db	715	ISQLRAQ	TGQTFFDYDPIITTTAPAAAAVVTAKQPEASLAKATSGIPASASAVSVAPAPT	774	RX	Kalman S., Mitchell W.P., Marathe R.W., Lammell C.J., Fan J., Hyman R.W.,				
Qy	486	SKLLKIN	-----DGEYTGDIYPA-----NGSSTLYQNV	514	RA	White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,				
Db	775	KTFLKIN	APDTQPEIQKVAEAAQSAVYNGKIVFSGEKLSEDAKNPLNATSVIHNDV	834	RA	Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,				
Qy	515	TTEQGRV	ILREKALSVNSLSOTGSLVMEAGSTWDF--VTP-----QPPQOPPA	562	RA	Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,				
Db	835	SLEAGTL	VSSGAGLLVDFTQKEGSLIVWDGTSIITNTVASEGLQSRSTPPSPKNAI	894	RT	McClarty G., Salzbegg S.L., Eisen J.A., Fraser C.M.;				
Qy	563	-----ANQL	-----ITLSNHLSSLLANNVTVN	587	RT	"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";				
Db	895	PVIRAVS	KVIASSLINLERADSGAGAVVPTIEESPDGSIITNLAVNLDLSLENGKVIT-	953	RL	Nucleic Acids Res. 28:1397-1406 (2000).				
Qy	588	PPTNPAQ	SHPAVIGSTAGSVTISGPIFFEDLDATADYRDVWLGSQKINVLKQLGT	647	RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
Db	954	-----LA	AGSGSVTLTGDLQFQDSSQSNFYDN-PLLNKNFTLFLDI---S	996	RC	STRAIN=TW-183;				
Qy	648	KPAN	-----APSDTLGNEMPKYGVQGSWKLAWDPNTANNPGYTLKATWTKTGVN	698	RA	Geng M.M., Schumacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,				
Db	997	APDAEKI	TEGFNIIPQAGTSSN-----LGQGGKWEVTEVDKSGKVSPEKK--WVSTGYI	1050	RA	Schneider S., Pohl T., Essig A., Marre R., Melchers K.;				
Qy	699	PGPERVAS	LPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDDRDLGQG	758	RT	"The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis.";				
Db	1051	PTANRAT	LVPSVWCSAIDMRAFNQNLVEVSTEGEDPHRGLWISGISNFFHKDSTKVQEG	1110	RL	Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.				
Qy	759	YRISGYS	LGANSY-FGSSMFLGATFEVFGRKDVVVCNSNHACIGSVYLSQQAALCG	817	RP	NUCLEOTIDE SEQUENCE OF 658-936.				
Db	1111	FRHISG	VYGVSTQPTISNKNVMDLAFQMLGSKDYRLADARSHVYAASHTTKCEK-LVN	1169	RC	STRAIN=CWL029 / VR1310;				
Qy	818	SYLFG	-----DAPIRASVGFNGHMKTSYTPFAESDVWVWNNCLAGE	859	RX	MEDLINE=20007584; PubMed=10539856;				
Db	1170	HYTFSKR	KGAILARKPEKSPIIFDA--QUSYSLSHNSMTTKTTPNPSRSSGNNHCVAGE	1227	RA	Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,				
Qy	860	IGAGLP	TVIT-PSKLYNLNLRPFVQAEFSYADHESF--TEEGQARAFKSGHLNLSVPV	916	RA	Madsen A.S., Knudsen K., Falk E., Birkelund S.;				
Db	1228	LSGYLP	PLVDHPA---IEELFPVKLHIVFVQEDDFKGTGGTENRFQSAHFVNSLPL	1284	RT	"Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity.";				
Qy	917	GVPFDRCS	THPNKYSMAAYICDAYTISGTETLLSHQETWTTDAFLHARGVGVVRS	976	CC	Am. Heart J. 138:S491-S495(1999).				
Db	1285	GVRFEKTKL	-NTYNIRLAYQPDYRDAPKSKVFLPSVHTAWSTGTATNLRSQAMILDGS	1342	CC	-I- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (potential).				
Qy	977	MYASLT	SNIEVYGHGRYEYRDASRGYLSAGSRVP	1012	CC	-I- SIMILARITY: Belongs to the PMP outer membrane protein family.				
Db	1343	DHHHLT	DNLEVFCHGAFELRGSSRNYNVDIGGRYKF	1378	CC	-I- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.				
Qy	977	MYASLT	SNIEVYGHGRYEYRDASRGYLSAGSRVP	1012	CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.				
Db	1343	DHHHLT	DNLEVFCHGAFELRGSSRNYNVDIGGRYKF	1378	CC	EMBL; AB001627; AAD18589.1; -; Genomic DNA.				
Qy	977	MYASLT	SNIEVYGHGRYEYRDASRGYLSAGSRVP	1012	CC	EMBL; AB002193; AAF38165.1; -; Genomic DNA.				
Db	1343	DHHHLT	DNLEVFCHGAFELRGSSRNYNVDIGGRYKF	1378	CC	EMBL; BA000008; BAA98653.1; -; Genomic DNA.				
Qy	977	MYASLT	SNIEVYGHGRYEYRDASRGYLSAGSRVP	1012	CC	EMBL; AB017158; AAF98393.1; -; Genomic DNA.				
Db	1343	DHHHLT	DNLEVFCHGAFELRGSSRNYNVDIGGRYKF	1378	CC	EMBL; AJ133034; CAB37067.1; -; Genomic DNA.				


```

DR PIR; B81591;
DR PIR; C72078; C72078.
DR PIR; C86546; C86546.
DR PHCI-2DPAGE; Q92898; -.
DR TIGR; CP0308; -.
DR InterPro; IPR005546; Auto transeptbeta.
DR InterPro; IPR006315; AutoTransporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; AutoTransporter; 1.
DR Pfam; PF02415; Chlam PMP; 7.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRfams; TIGR01414; autotrans barL; 1.
DR TIGRfams; TIGR01376; POMP_repeat; 5.
DR Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
KW SIGNAL.
FT CHAIN 24 936
FT CONFLICT 658 666 PTHGFRHI -> EDNIRYRN (in Ref. 5).
FT CONFLICT 822 822 Y -> H (in Ref. 1, 4 and 5).
SQ SEQUENCE 936 AA; 100107 MW; 3981DB3C950AP95A CRC64;

Query Match 19.7%; Score 1040; DB 1; Length 936;
Best Local Similarity 29.0%; Pred. No. 7.8e-53;
Matches 310; Conservative 162; Mismatches 406; Indels 192; Gaps 31;

QY 1 MQTSPHKFPLSMILAYSCSLNGGYYAAHIV--PQGIYDGETLTVSPYTVIGDPGOTT 58
DB 1 MKSVSWLFFS8IPLPFSLSI-----VAAEVTLDSNNSYDGSNGTTFVFTTDAAGTT 56

QY 59 VFSAGELTKNLDNSIAALPLSCFNLLGSPVTLGRGHSHTFENIRT-STNGAALSNSAA 117
DB 57 Y8LLSDVFNQAGALGIPLAGCFLEAGCDITFQGNQHALPAPINAGS8AGTVASTSAA 116

QY 118 DGLFTIEGKELSFNCNSLLAIVLPAATNNKGSQPTTTTSPNGIYKSTDLNNEK 177
DB 117 DKNLLFNDFSLSIIFCSPLLL-----SPT-----GQCALXSVGNLSLTGNSQ 159

QY 178 FSPYENLVSGDGAIDAKSLTVQIGISKLCVQENTQAQDGGACQVTVSFSMAEAPIAF 237
DB 160 TIFTQNFSDNGGVINTKNFLSCTSPASFNRQA----- 195

QY 238 VANVAGVGGGIAAQDQCGQVSSSTSTEDP-VVSPSRNTAEPDGNVARGVGGIYSXGN 296
DB 196 ---FTGKGQGVYA-----TGTITIENSQIVSFQNLK-KGS-----GGALYSTDN 238

QY 297 VAPLNGKTLFLNNVAPVYIAAKOPTSGQASNTNNYDGGAIKCKGAKGAGNNSGSV 356
DB 239 CSITDNFQVIFDQNSA---WEAAQ-----AQQGAICC-----ITTDKTV 274

QY 357 SPDGEVGVFSSNVAGKGGAIYAKLKVANGCPVQFLRNTAND-----GGAIYLGESG 410
DB 275 TLTKGNKLSPTNTALTATYGAISGLKVK8ISAGGPTLFQSNISGSGSAGQGGGAINIASAG 334

QY 411 ELSLSADYGDIIIPQNLKRTAKENAADVNTVSSQAI5SMGSGGKITTLLRAKAGHQLLEN 470
DB 335 ELALSATSQDITFNN-----QVTNGSTSTRNALNIIDTAKVTSIRATQOSIIFY 385

QY 471 DPEIEMANNQPOASSKLLKINDGEG-----YTGDIVF-----ANGSSTLYQ 512
DB 386 DPIT-----NPGTAASDTLNLNLADANSEIYEGGAIYVSGKLSPTKAIKAAVNTSIRQ 441

QY 513 NVTIEQGRIVREKAKLSVNSLSQTGGS-LYWEAGSTWDFVTPPPQPPPAANQLITLNS 571
DB 442 PAVLARGDLVLDGVTVTFKDLTQSPGSRILMDGGTLLS-----AKENLSLNG 490

QY 572 LHLSSLSLANNVNTPTNPQDSHPAVIGSTTAGSVTISGPIFFEDLDLDYDQVDM 631
DB 491 LAVNLSSLDGNTKA-----ALKTEAADKNISLSGTALIDTSGSYFENN- 535

QY 632 LGSNQKINVLKQLGTKEPPAN-----APSDLTILGNEMPKYGYQGSWKLANDPNTANGP 685
DB 536 LKSASTYELLEL---TTAGAGNTITLGAISLTLLQEPETHYGYQGNMQLSWANATSSK-- 590

```

RESULT 15

PMP9 CHLPN

ID -PMP9 CHLPN STANDARD; PRT; 928 AA.

AC Q92398;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Probable outer membrane protein pmp9 precursor (Polymorphic membrane

protein 9) (Outer membrane protein 10).

GN Name=pmp9; Synonyms=omp10; OrderedLocusNames=Cp00447, CP0306, Cp0464;

OS Chlamydia pneumoniae (Chlamydia pneumoniae).

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OC NCBI_TaxID=83558;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=CWL029 / VR1310;

RC MEDLINE=20007584; PubMed=10539856;

RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,

Madsen A.S., Knudsen K., Falk E., Birkelund S.;

RT "Molecular biology of Chlamydia pneumoniae surface proteins and their

role in immunopathogenicity.";

RL Am. Heart J. 138:S491-S495(1999).

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=CWL029;

RC MEDLINE=20150255; PubMed=10192388; DOI=10.1038/7716;

RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,

Olinger L., Grimwood J., Davis R.W., Stephens R.S.;

RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";

RL Nat. Genet. 21:385-389(1999).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=AR39;

RC MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.P.,

White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,

Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,

Dodson R.J., Winn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,

RA McClarty G., Salzberg S.L., Bisen J.A., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia

pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

RN [4]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=J138;

RC MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi P., Ouchi K.,

Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;

"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.",
Nucleic Acids Res. 28:2311-2314 (2000).
[5]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Warre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -!- SIMILARITY: Contains 1 autotransporter (TC 1.B.12) domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: AJ133034; CAB37069.1; -; Genomic DNA.
CC EMBL: A2001628; AAD18591.1; -; Genomic DNA.
CC EMBL: A2002192; AAP38163.1; -; Genomic DNA.
CC EMBL: BA000008; BAA98655.1; -; Genomic DNA.
CC EMBL: A2017158; AAP98395.1; -; Genomic DNA.
CC PIR: B72077; B72077.
CC PIR: B86546; B86546.
CC TIGR: CP0306; -.
CC InterPro: IPR005546; Auto transportbeta.
CC InterPro: IPR006315; Autotransporter.
CC InterPro: IPR011427; ChlamPMP_M.
CC InterPro: IPR003368; Chlamydia_PMP.
CC InterPro: IPR006626; PbhA.
CC Pfam: PF03797; Autotransporter; 1.
CC Pfam: PF02415; ChlamPMP; 7.
CC Pfam: PF07548; ChlamPMP_M; 1.
CC SMART: SM00710; PbhA; 5.
CC TIGRFAMs: TIGR01414; autotrans bar1; 1.
CC TIGRFAMs: TIGR01376; POMP repeat; 6.
CC Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 928 Probable outer membrane protein pmp9.
SQ SEQUENCE 928 AA; 98333 MW; 58910A8F04F12219 CRC64;

Query Match 19.4%; Score 1021; DB 1; Length 928;
Best Local Similarity 29.5%; Pred. No. 1e-51;
Matches 313; Conservative 161; Mismatches 404; Indels 184; Gaps 39;

QY 1 MOTSFHFKPLSMILAYSCCSLNGGVYAEIMV---PQGIYDGE-TLTVSPYTVIGDPDG 56
DB 1 MKSSLHFWLFISSSLALP-LSLNFSAAPAARVVEINLPTNSPFGPTVT---PPAQTTNADG 56

QY 57 TTVFSAGELTKLNDLSIALPLSCFNLGLSPVLGRHSLTFENRTSTNGAALSNSA 116
DB 57 TIYNLTGDVSIITNA-GSFALTALTAKTBTGNSLFGQHGQYQLQNDAGAN-CTFTNTA 114

QY 117 ADGLFTIEGFKELSFNSCNLSLAVLPAATNKGSTPTTTSTPENGTYIKTDLILLNNE 176
DB 115 ANKLLSPGSGSYLSL-----IQTWATTTGTGAKTGCAGSIQSNY 154

QY 177 KFSFYSNLVSGDGAIDAKSLITVQGISKLCVQENTAQADGGACQVVTFSAMANEAPIA 236
DB 155 SCYFGQNFNDGALQSSISL-SLNPNTLFAKNKATQKGGAL-----YST----- 200

QY 237 FVANVAGVGGIAAVQDGGQGVSSSTSTEDPVVPSRNTAVEFDGNVARGGIIYSGN 296
DB 201 -----GGIT-----INNTL-----ASFSENTA-----ANNNGGAIYTEAS 231

QY 297 VAFLLNGKTL-FLANVASPVYIAKAQPTSGQASNTSNNGYDGGAIKFCNGAQAQSNNSGS 355
DB 232 -SFISNNKAISFINN-----SVTATSAT-----GGAIYC-----SSTSAPKPV 268

QY 356 VSFDEGVVFPSSNVAAGKGAIYAKKLSVANCGPVPQFLRNAND-----GGAIYLGSSG 410
DB 269 LTLSDNGELNFIGNTAITSGGAIYTDNLVLSGGGPTLPKNSAIDTAAPLGGAIADSG 328

QY 411 ELSLSADYVDIIFDGNLKRKTAKENADVGVTVSSQAISMG-SGGKITTLAKAKGHQILP 469
DB 329 SLSLSALGDITFEGN---TVKGGAS---SSQTTTRNSINIGNTNAKIVOLRASQGNITLYP 383

QY 470 NDPIEMANGNQPQSSKLLKINDGE-----GYTGDIVPA-----NGSSTLY 511
DB 384 YDPIT-----TSITALSALMLNGLPDLAGNPAYQGTIVFSGEKLSEAEADNLKSTIQ 439

QY 512 QNVTTIEQRIVLREKAKLSVNSLSQTGS-LYMEAGSTWDFVTPOPPQPPAANLILTS 570
DB 440 QPLTLAGQLSLKSGVTIVAKSFQSPGSTLLMDAGTTLETADG-----ITIN 487

QY 571 NUHLSLSLLANNVNTPTPPAQDSHPANVIGSTTAG-SVTISGPIPEDDLDDTDAYDRY 629
DB 488 NLVLNVDLSL-----KETKATLKATQASQVTVTLGSLSLVDPSPGNVYEDV 532

QY 630 DWLGSNOKINVLKLOLGTKPPANAP-SDLT---LGNEMPKYGYQGSWKLWDPNTANNGP 685
DB 533 SW--NNPQV-FSCLTLTADDDPANIHIITLADPLKPNPHWGYPQGNWALSQWQEDATYKSK 589

QY 686 YTLKATWTGTGYNPGPVRVASLVPNSLMGSIIDIRSAHSAIQASVDGRSYCRGLWVSGVS 745
DB 590 -NATLTWTGTGYNPNPERRGTLVANTLMGSPVDVRSIQQLVATKVRQSQETRGWCEGIS 648

QY 746 NFFYHRRDALGQGYRISGGYISLGANSYFGS-SMFGLAFTFVGRSKDYVVCRSNHHACI 804
DB 649 NFFHKDSTKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKDRDHFINKRASAYA 708

QY 805 GSVYLSLTOALCG-----SVLFG-----DAFIRASYGFGNOMKTSYTPAEESDVRW 851
DB 709 ASLHLQHLATLSSPRLRYLPGSESEQVLFDAQI---SVIYSKNTWKTYYTOAPKGSSEW 766

QY 852 DNNCLAGIGAGLPITVITPSKLYLNELPFVQAEFSYADHESFTEGDC-ARAPKSGHLL 910
DB 767 YNDGCALELASSLPHTALSHGLFPHAYFPPIKVEASYIHQDSFKERNNTTLVRSFDSGLI 826

QY 911 NLSVPVGVKDFRCSSHTPHNKYSFMAAICDARTISGTSTTLLSHOETWTTDAPHARHG 970
DB 827 NVSPVIGITFERSRNERASYEATVIYADVVRKPNPDCCTALLINNTSWKTTGTNLSRQA 886

QY 971 VVVRGSNYASLNTSEVYGHGRYEYEDASRGYGLSAGSRVP 1012
DB 887 GIGRAGIFYAFSPNLEVTNLSMEINGSSRSYNADIGGRKQF 928

RESULT 16
PMP6_CHLPN STANDARD; PRT; 1276 AA.
ID AC Q9Z899; Q9JRW2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable outer membrane protein pmp6 precursor (Polymorphic membrane
protein 6).
DE Name=pmp6; OrderedLocusNames=CPn0444, CP0309, CPn0460;
GN Chlamydia pneumoniae (Chlamydophila pneumoniae).
OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OC NCBI_TaxID=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].


```
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shira M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and W0209 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TW-183;
RA Geng M.M., Schummacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BA000008; BAA98226.1; -; Genomic DNA.
DR EMBL; AE017157; AAP97951.1; -; Genomic DNA.
DR PIR; H86492; H86492.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Auto transporter.
DR InterPro; IPR005546; Auto transportbeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR TIGRFAMs; TIGR01414; autotrans barl; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
SQ SEQUENCE 772 AA; 82931 MW; 484FC56D635801EB CRC64;

Query Match 18.6%; Score 978.5; DB 2; Length 772;
Best Local Similarity 29.4%; Pred. No. 2.6e-49;
Matches 260; Conservative 142; Mismatches 316; Indels 167; Gaps 26;

QY 180 FVSLVSGDGAIDAKSLTVQIGSLKLVFOENTAGDGGACQVVTSPSAMANAEPAPVA 239
DB 3 FPKFSTDNGAINTAKTSLTGTSALFSENSTSKGGAIGTSDALTITGNQGEVSP-- 60

QY 240 NVAGVRGGIAAVDQGGVSSSTSTEDPVVFSRNTAVEFDGNVARVGGGIYSYGNVAF 299
DB 61 -----SDNTSSDS-----GAAIFTEASVTI 80

QY 300 LNNKTLFLANNVSPVYIAAQPTSGQASNTSNNGYDGGALFCKNGAAGNNSGSVSFD 359
DB 81 SNNKVSFDN-----KVTGASSSTGDM--SGGAI CAYK-----TSTDTKVTLT 123

QY 360 GEGVFFSNAAGKGAIAKLSVANGCPVQPLRNIA-----DGGAIYLGSGGSL 414
DB 124 GNQMLLFNNSTTAGGAIYVKLELSSGGLTLFSRNSVNGTAPKGAIAIEDSGELSL 183

QY 415 SADYGDIIFDGNLKRITAKENADVNGVTVSSQAISSMGSGGKITTLRAKAGHQIILFNDPIE 474
DB 184 SADSGDIVLGN---TVSTTTPGNT-----RSSIDLGTSAKWTALRSAGRAIYFDPI 235

QY 475 MANGNNQPAQSKLLKINDGEG-----YTGDIVFA-----NGSSTLYQNVTI 516
DB 236 TGSSTT-----VTDVLKNVETPADSALQYTGNIIFTEKLSLSEARDSKNLTKLQPVTL 291

QY 517 EQGRVILREKAKLSVNSLSQTGG--LYMEAGSTWDFVTPQPPQPPAANQLITLSNLHLS 575
DB 292 SGGTSLSLKHGVTLTQATQADRLENMDVGTLE-----PADTS---TINNLVIN 339

QY 576 LSSLANNNAVNTPPNPPADSHPAVIGS--TAGSVTISGPIFFEDLDOTAYDRYDMLGS 634
DB 340 ISSI-----DGAKKAKIETKATSKNLTLSTGVTITLDDPTGTFYENHS--LRN 383

QY 635 NQKINVLKQL--GYKPPANASDLTLGNEMPKYQGSW--KLAWDPNTANNGPPTLKATW 692
DB 384 PQSYDILELKASGVTVSTAVTDPDPMGEKF--HYGQGTMGPIVNGTGAISTATP-----NW 438

QY 693 TKTGNGPGRVASLVPNSLWGSTILDIRSAHSALQASVDGRSYCKRGLWVSGVSNFFYHDR 752
DB 439 TKTGYINPEIGSLVPSNLNNAIDISLHYLMETANEGLQGDRAFCAGLSNFFHKDS 498
```

```
QY 753 DALCGQVRYISGGYSLGANSYFGS--SMFGLAFTVEFGRSKDYVVCRSNHHACISGVYLS 811
DB 499 TKTRGRFHLISGGYVIGGNLHSTCDKILSAAPCOLFGRDRDYFAKNG-----GTVYGGT 553
QY 812 QOALCGSVLFGDAFI-----RASYGFGNQHMKTSTYFAES 847
DB 554 LY-----YQNETYISLPCKLRPCSLSYVPTFIPVLSGNLSYTHTDNDLTKTKYTYPTV 608
QY 848 DVRDNNCLAGEIGAGLPVITPSKLYLNEIRPFVQAEFSYADHESFTTEGDQARAPKSG 907
DB 609 KGSNGNSPALEFGGRAPICLDESALP--EQWMPKMLQFVYAHQEGKEQTEAREFGSS 667
QY 908 HLNLNLVPGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFLA 967
DB 668 RLNLALPIGIRPKESDCQDATYNTLGYTVDLVRSNPDCTTTTLRISGDSMKTFTGTLA 727
QY 968 RHGVVVRGSMYASLTNSIEVYGHGRYERDASRGYGLSAGSRVPF 1012
DB 728 QOALVLRAGNHFCFNSNFEAFSPQSFELRGSSRNYNVDLGAKYQF 772

RESULT 19
Q82205 CHLCV
ID Q82205_CHLCV PRELIMINARY; PRT; 841 AA.
AC Q82205;
DT 01-JUN-2003 (TRMBLrel. 24, Created)
DT 01-JUN-2003 (TRMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family
DE protein/autotransporter.
GN OrderedLocusNames=CCA00624;
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaaceae; Chlamydiaophila.
OX NCBI_Taxid=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapple E.K., Khouli H.M., Pederoza N.B.,
RA Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaaceae.";
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016996; AAP05366.1; -; Genomic DNA.
DR TIGR; CCA00624; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Auto transporter.
DR InterPro; IPR005546; Auto transportbeta.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR TIGRFAMs; TIGR01414; autotrans barl; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 3.
KW Complete proteome.
SQ SEQUENCE 841 AA; 90411 MW; 21A9ED16E0D7C65D CRC64;
```

```
Query Match 18.4%; Score 971.5; DB 2; Length 841;
Best Local Similarity 27.7%; Pred. No. 7.6e-49;
Matches 298; Conservative 132; Mismatches 347; Indels 299; Gaps 30;

QY 1 MQTSFHKKFFLSMILAYSCCSLNGGGYAA---EIMVPOGIYDGETLTYSF-----PY 48
DB 1 MKHPVYVFLVSSGLLASTSS---SPAAAVOETLNSDSSYNGNTTTTTFVFPKSTGAEY 56
QY 49 TVTGDPDSGTTFVSAGELTLKLNLSIAALPLSCFGLNGSFTVLGRGHSITPENIRTSN 108
```

```
Db 57 TCNGN---VCITYAGKTT-----PLTKSCFTETTENLFLGQGYSLCFDNIATAK 104
Qy 109 GAALNSAADGLFTTEGKELSFNCNSLLAVLPAATTKGSGQTPTTSTPSNGLIYKT 168
Db 105 PAATIEVSANDKTLISGFSFLPSCSYC-----PPGTTEQ 137
Qy 169 DLLLNNKFPSPYSLNLSGDGAIDAKSLTVQGISKLCVFOENTAQADGACQVVTSPSA 228
Db 138 -----GAIQSGVT----- 146
Qy 229 MANEAPIAFVANVAGVRGGIAAQQGQGVSSSTSTEDPVVFSRNTAVEPDGNVARVG 288
Db 147 ----- 146
Qy 289 GGIYSYGNVAFPLNGKTLFLNNVAVPVYIAAKOPTSGQASNTSNVYDGGAIFFCKNGQA 348
Db 147 -----TFGNNDKIIFPKNCSTE-----KGAIKCDTGTNA 176
Qy 349 GSNNSGSVDFGEGVFPFSSNVAAGKGAIAKCLSVANGCPVQFLRNIA-----DGCA 403
Db 177 -----ELKPEGNKYLFPSCNSQOEGGAIYAKLSIISGGTLPFSNNTSKAADPKGA 230
Qy 404 IYLGHS-GBLSLSADYGDIIIPGNLKRATAKNAADVGVTVSSQAISMGSGGKITTLRAK 462
Db 231 ICIAADADSECSLTAENGDIIPGNKIITGTPTKEN-----SIDLGGGKFSQLRAR 283
Qy 463 AGHQILFNDPIEMANGNPOASSKLLKINDGEG---YTGDIIVPA-----NG 506
Db 284 DGFQGVFFYDPI-----ANN--GSDTDTLEINKADGAATYSGRIVFSGEKLTDEKQVTDNL 337
Qy 507 SSTLQNVITIEGRVILREKAKLSVNSLSGTGSLVMEAGSTWDFVTPPQQPPQPAANQL 566
Db 338 KSFFQKPLTVGSGFVLKNGVTVSAKQITQSGAIEADAGTN----- 379
Qy 567 ITLSNLHLSSLANNAVTPNPTNPPAQDSHPAVIGSTAGSVTISGPIFFEDLDITAY 626
Db 380 LSTTIEDISLNLVINTASIGGGVPLAAQ-----ISAEGTKSVTIS--LNLVDAGNGY 434
Qy 627 DRYDMLGSNQKINVLKQL-GTKPPANASDLTLGNEMPKYGVQSGKWLAWDNTANNGP 685
Db 435 EYVPFSTTRFPFSLIEAKANGTGTPTIPTTLTHDHAFAHYGYGLWTTSWAQGTATTSQ 494
Qy 686 YILKATWTKTGNPGERVASLVPNSLWSILDIRSAHSAIOASVDGRSVCRGLWVGVS 745
Db 495 LATLA-WQOTGYNPNPERQGPLVNTLWGSFSDVRAIQNLMDISVNGADYQRLWASGLA 553
Qy 746 NPFYHRRDALGQYRIVISGYSLGANS-YFGSMFGLAFTEVGRSKDYVVCRSNHACI 804
Db 554 NFIQKSGTETKFRHHSAGVVLGAYAKTLLSDVFSAAFCOLFGRDKDVLVSQNSNIYA 613
Qy 805 GSYLSTQQALCGSYLFGDAP-----TRASYGFGNQHKMTSYT--FAE 845
Db 614 GSIYYQHTS-----FWDADWNLQSTLGAQAPLVLAQLTYSHTSNDKMTNMTTKYAP 666
Qy 846 ESDV-----RWNQNLGHBIGAGLIVITPFSKLYLNLRLPFPVQAQSFYSADHESPTTEGD 899
Db 667 QGVYYPEIKGDWNGDCFGVELGATVPIESPSLFP-DMYSFPLRFQLVVAHQSDFKENNS 725
Qy 900 -QARAFKSHLLMLVVPVGVKFCRCSSTHBNKYSFMAAYICDAYRTISGTETTLTSHQET 958
Db 726 TEGRYFESSDLTWSLWPIGVKFRPFSNDNDIASVNTLAVAPDLVRSNPDCKTSLVSPPT 785
Qy 959 --WTTDAFLHARGVVRVGSVMYASLTISNIEVYGHGRYBYRDAAGYGLSAGSRVF 1012
Db 786 AVMLTKATNLARHAFIVKAGNYLSLSSNPFISQFGFELGSSRTYNVDLGSKIQF 841

RESULT 20
PMP1 CHLPN
ID PMP1 CHLPN STANDARD; PRT; 922 AA.
AC Q929G5; Q9KLY9; Q924H9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
```

```
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable outer membrane protein pmp1 precursor (Polymorphic membrane
DE protein 1) (Outer membrane protein 6).
GN Name=pmp1; Synonyms=omp6; OrderedLocusNames=CpN0005, CP0770, CpB0006;
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae (Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CWL029 / VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christensen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
RN (2)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN (3)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey R.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher E.K., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN (4)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10671362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN (5)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RX Geng M.M., Schumacher A., Muehlhofer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
CC Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ133035; CAB37075.1; -; Genomic DNA.
CC EMBL; AB001585; AAD18163.1; -; Genomic DNA.
CC EMBL; AB002237; AAF38570.1; -; Genomic DNA.
CC EMBL; BA000008; BAA98215.1; -; Genomic DNA.
CC EMBL; AB017157; AAP97939.1; -; Genomic DNA.
CC FIR; B72131; B72131.
CC FIR; B86491; E86491.
CC TIGR; CP0770; -.
CC InterPro; IPR005546; Auto transptbeta.
CC InterPro; IPR011427; ChlamPMP_M.
```


Query Match		18.0%; Score 950; DB 2; Length 847;
Best Local Similarity		27.4%; Pred. No. 1.5e-47;
Matches 282; Conservative 145; Mismatches 332; Indels 270; Gaps 33;		
QY	37	YDGETLTVSPFVTVIGDPGGTTFVSAGELTL--KNLDSIAALPLSCFGLGSLGFTVLGR 94
DB	36	YNGNTNSRFPNPLSTNSNGTIYCTGNICIAVAGLDGS--GLSSSCFTDTAGNLSFLGN 93
QY	95	GHSHTFENIRT-STNGAALSNAGADGLFTIEGFKELSPNCNSLLAVLPAANTNKGSTQP 153
DB	94	GVTLCFDNITQSSHGPAISVGTNKTLDISGSLFSCAYC-----134
QY	154	TTTSPNSNTIYKTDLLLLNNEKPFYENLVSGDGAIDAKSLTVQGISKLKCVQDENTA 213
DB	135	-----PPGATGY-----141
QY	214	QADGGACQVTVSPSAMANEAPAFVANVAGVRGGIAAVQDGGQGVSSSTSTEDPVVVS 273
DB	142	-----141
QY	274	RNTAVEFDGNARVGGIYSYGNVAFNNGKTLFLNNAVSPVIAAKQPTSGQASNTSN 333
DB	142	-----GAIRAVGNTTIKONSSLVFHKNCST-----166
QY	334	YDGGAIKFCNGAQAQSNNGSVSPDGGVVPFSSNVAAKGGKGAIVAKKLSVANCGPVQF 393
DB	167	-GEGGAIQCK-----ASSSEAEKLIENNOQLVPAENSSSSSGAIIYADKLTIVSGGPTLF 220
QY	394	LRNIAN-----DGAIVLGHES-GBLSLSADYGDIIIFDGNLAKTKAKNAADVGVTVSSQAI 448
DB	221	SNNSVASSPKGAICIKSGGSCSLTADLGDITFDGN--KIKTNG--GSPVTVTNSI 275
QY	449	SMGSGGKITTLRAKAGHQLFNPDPIEMANGNNOPOASSKLLKIN--DGEYTGDIYVF--- 503
DB	276	DLGSSGKFTKLNKAGEGFIFFYDPI-TGGGSDE-----LNINKQDVTVDYTKIVFSGE 327
QY	504	-----ANGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSOT-CGSLYWEAGSTWDFV 552
DB	328	RLSDEBKVAANLKDSPKQPKLGSGSLIILKDGVTLETSTFTQTEGATVMDLGTTL--- 384
QY	553	TPQPPQPPAANQLITLSNLSLSLANNVNPPTNPPADDSHPAVIGSTTAGSVTI 612
DB	385	-----QTPSSGGETITLNDIINVASLGGGVADPP-----AKVEATTESKVTI 429
QY	613	SGPIFFEDLDDTAYDRYDNLGSGNQKINVLKQLG-----TKPPANAPSDTLGNEMP--K 665
DB	430	NAVNLVDD-NGNAYE-YPILAASQPPTAIEVRSGSGSIITKPTTN-----LENYTPPTH 481
QY	666	YGYGSKLAWDPNTANGPYTLKATWTKTGYNPGPERVASLVPNSLWGLSILDIRSAHSA 725
DB	482	YGYQGNVTWTKQSSAQ-EKATATLTWEQTGYSPNPERQGSLENTLWGSFSDIRATQNL 540
QY	726	IQASVDGRSICRLWVSGVSNFFVHDDRALQGGVRYISGGVSLGANSYFSG-SMFGLAFT 784
DB	541	MDISVNGADYHRGFWPWSGLNFLHKSQSDTKRKRHNSAGVALGVYAQTTPSEDVFSAPFC 600
QY	785	EVFGSKDYVYCRNHHACIGSVYL-----STQALCGSYLFGDAFI-----RASYGFGN 834
DB	601	QLFGKDKDLYVKNSSVTYAGSIYYQHSYWNVTWNTLLQNTLGAELPLVLAQLAYCHAS 660
QY	835	QHKMYSY--TPAEB-----SDVR--WDNCLAGETGAGLPVITPSTKLYLNELRFPVQABF 886
DB	661	NNKTNMTDTPAPKTTYSEIKDQWGNDCFCVBFCAKAP--IETASLLFDMYSPFVKQL 718
QY	887	SYADHSEFTE-EGDQARAFKSGHLLNLSVPVGVKFDRCSSTHPNKYSFMAIYICDAVRTI 945
DB	719	VHAHQDDPKENNSDQGRYFESNNITNLSMPIGVKLEKPSHKDTSYNTLTAIAPDIYRSN 778
QY	946	SGTETTL--SHQETWTTDAPLARHGVVWAGSVYASLTSNIERYVGHGRYRDSRGYG 1003
DB	779	PDCTASLLVSPTSVAWVTKANNLARHAFILQAGNYLALTRNTELFSPQGFELGSCRTYN 838

QY 1004 LSAGSRVRF 1012
DB 839 IDLGSKIQF 847

RESULT 22

P77792	CHLAB
ID	P77792-CHLAB PRELIMINARY; PRT; 839 AA.
AC	P77792; Q5L5P5;
DT	01-FEB-1997 (T-EMBLrel. 02, Created)
DT	01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
DT	13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE	POMP90B precursor (POMP90A precursor) (Polymorphic outer membrane protein).
GN	Name=pmp17G; Synonyms=pomp90B; OrderedLocusNames=CAB598;
OS	Chlamydomonadales; Chlamydiales; Chlamydiales; Chlamydiales.
OC	Bacteria; Chlamydiales; Chlamydiales; Chlamydiales.
OX	NCBI_TaxID=83555;
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=S26/3;
RC	MEDLINE=98187897; PubMed=9529048;
RA	Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT	"Molecular cloning and characterization of the genes coding for the highly immunogenic cluster of 90-kilodalton envelope proteins from the Chlamydia psittaci subtype that causes abortion in sheep.";
RL	Infect. Immun. 66:1317-1324 (1998).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=S26/3;
RC	MEDLINE=96406378; PubMed=8810511; DOI=10.1016/0378-1097(96)00281-9;
RA	Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RT	"Identification of a multigene family coding for the 90 kDa proteins of the ovine abortion subtype of Chlamydia psittaci.";
RL	FEMS Microbiol. Lett. 142:277-281 (1996).
RN	[3]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC	STRAIN=S26/3;
RC	PubMed=15837807; DOI=10.1101/gr.3684805;
RA	Thomson N.R., Yeates C., Bell K., Holden M.T.G., Bentley S.D.,
RA	Livingstone M., Corden-Farrago A.-M., Harris B., Doggett J.,
RA	Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
RA	Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;
RT	"The Chlamydomonadales abortus genome sequence reveals an array of variable proteins that contribute to interspecies variation.";
RL	Genome Res. 15:629-640 (2005).
DR	EMBL; U65943; AAC15924.1; -; Genomic DNA.
DR	EMBL; U65942; AAC15922.1; -; Genomic DNA.
DR	EMBL; CR848038; CAF64045.1; -; Genomic DNA.
DR	GO; GO:0019867; C:outer membrane; IEA.
DR	InterPro; IPR005546; Auto transporter.
DR	InterPro; IPR006315; Auto transporter.
DR	InterPro; IPR011427; ChlamPMP M.
DR	InterPro; IPR003368; Chlamydia PMP.
DR	Pfam; PF03797; Autotransporter; 1.
DR	Pfam; PF02415; Chlam PMP; 4.
DR	Pfam; PF07548; ChlamPMP M; 1.
DR	TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR	TIGRFAMs; TIGR01376; POMP repeat; 3.
KW	Complete proteome; Signal.
FT	SIGNAL 1 16 Potential.
FT	CHAIN 17 839 POMP90B.
SQ	SEQUENCE 839 AA; 89825 MW; 4581C7CBAP7FP4C4 CRC64;
Query Match 18.0%; Score 949.5; DB 2; Length 839;	
Best Local Similarity 27.3%; Pred. No. 1.5e-47;	
Matches 291; Conservative 153; Mismatches 341; Indels 281; Gaps 36;	
QY	1 MQTSFKHFPFLSMILAYSCCSLNGGGVAAETWVQGIYDGETLTVSPFPTVIGDPSTVTF 60
DB	1 MKHPVYWFLLSSSL-FASNSLSFANDQATLTALTPSDSYNGNVTSEEP--QVKETSSGGTYYT 57
OY	61 SAGELTLKNLDSIAALPLSCFGLGSLGFTVLGRHSLTPTENTRTSTNGAALSNAADG- 119

```
Db      58 CEGNVC1-SPAKDSGLKSCF-SATDNLTLFGNGYTCFNDITNTASPGAINVOGGK 115
Qy      120 LFTIEGFKELSPNCSNLAVLPAATNNGS-QTPTTSTPSNGYTSKTDLLLNNEKF 178
Db      116 TLIGSPSLFSCAYC-----PPGTGGAIOGTGNTTLKDNSSL----- 154
Qy      179 SPYSLVSGDGAIDAKSLTVQGISLKVQFQNTAQAGGACQVTSFSMANEAPIAFV 238
Db      155 -----VFHKNCSTAEAGAIQ----- 169
Qy      239 ANVAGVGGGIAAVQDQGGVSSSTEDPVVPSFENTAVEFDGNVAVGVGGIYSGNVA 298
Db      170 ----- 169
Qy      299 FLNNGKTLFLNNAVSPVYIAAKQPTSGOASNTSNNGYDGGAIKCKNGAQAQ-----SNNSGS 355
Db      170 -----CKGSSDAELKIENNQN 186
Qy      356 VSPDEGVVFPSSNVAAGKGAIIYAKLSVANGCPVQFLNIAND-----CGAIYLG-S 409
Db      187 V-----FSENSSTSGGAIYADKLTIVSGGPTLFSNNSVSGSPKGAISIKDSS 237
Qy      410 GELSADYGDIIIPDN-LKRTAKENAADVNGVTVSSQAIWSGSGKITTILRAKAGHOIL 468
Db      238 GECSTADLGDITFGNKIKITSGSS-----TVRNSIDLGT-GKFTKLAKDGGFIF 290
Qy      469 FNDPIEMANGNNQPAQSSKLKINDGE--GYTGDIVPA-----NGSSTLYQN 513
Db      291 FVDPI-TGGGDE-----LNINKETVDTGKIVFSGEKLSDDEKARAENLASTFNQ 342
Qy      514 VTIEGRIVLREKAKLSVNSLSQDGS-LYWEAGSTWDFVTPQPPQPPAANQILITSLN 572
Db      343 ITLSAGSLVLDGVSVTAKQVTOEAGSTVMDLGTTL-----QTSSGGETITLTNL 394
Qy      573 HLSLSLLANNAVTPPTPPAQDSHPAVIGSTTAGSVTISGPIFFEDLDTDVDRYDML 632
Db      395 DINTASLGGGGTS-----PAKLATNTASQAITTNANVLVDADGNAYE--DPI 440
Qy      633 GSNQKINVLKLGQTKPPANAPSDITLGNEMP--KYGYQSGSKLAWDPNTANNQPYTLKA 690
Db      441 LATSKPPTAIVATTWASTVTPQTD-NLNYVPPTHYGYQGNWTVTMDTETATK---TATL 496
Qy      691 TWTGTGYNPGERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGLWGVSGVSNPFYH 750
Db      497 TWEQTGYSPNERQGPLVNTLWGAFLDRAIQLMDISVNGADYHRRGFWYSGLANFLHK 556
Qy      751 DRDALGCGYRIVISGYSILGANSYFGS-SMPGLAPTEVFGSKDYVVCNSHHCIGSVYL 809
Db      557 SGSDTKFRNSAGYALGVYAKTPSDDDIFSAAPCQLFGKDKDLVLSKNANIYAGSLY 616
Qy      810 -----STQQAALCGSYLFGDAFI-----RASYGFGNQHKTSY-----TFAE-BSDV 849
Db      617 QHISVSAWNLQNTICABAPLVLAQLTYCHASNQDKTNWTTYPAPKTYTAEIKGD- 675
Qy      850 RWDNCLAGETGAGLPVITPSKLYNELRFPVQAFSADHESFTE-EGQOARAPKSGH 908
Db      676 -WGNDCFGVELGATVPIQ-TESSLIFDMYSPLKFLVHTHQDDPKENNSDQGRYFSSN 733
Qy      909 LLNLSPVGVKFRDCSSTHPNKYSFMAAYICDAVRTISGTETTLTSHOET--WTTDAFHL 966
Db      734 LTNLSLPIGIFERFANNNDTASYHVTAAYSPDIVRSNPDCTTSLVSPDSAVVTKANNL 793
Qy      967 ARHGVVVRGSMYASITNSIRYVHGRIYRDRASRGYGLSAGSRVRF 1012
Db      794 ARSAPMLQAGNYSLSHNIETIPSGFGLGSSRTYVNDLGSKIQF 839
```

RESULT 23

PMP2_CHLPN

ID_PMP2

AC_Q923A1, Q9RB73;

DT 16-OCT-2001 (Rel. 40, Created)

STANDARD; PRT; 841 AA.

```
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-MAY-2005 (Rel. 47, Last annotation update)
DE      Probable outer membrane protein pmp2 precursor (Polymorphic membrane
DE      protein 2) (Outer membrane protein 7).
GN      Name=pmp2; Synonyms=omp7;
GN      OrderedLocustNames=CpN0013, CpB0015/CpB0016;
OS      Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX      NCBI_TaxID=83558;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CWL029 / VRL310;
RC      MEDLINE=20007584; PubMed=10539856;
RA      Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA      Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT      "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT      role in immunopathogenicity.";
RL      Am. Heart J. 138:8491-8495(1999).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=CWL029;
RC      MEDLINE=99205606; PubMed=10192388; DOI=10.1038/7716;
RA      Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Pan J., Hyman R.W.,
RA      Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT      "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL      Nat. Genet. 21:385-389(1999).
RN      [3]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=AR39;
RC      MEDLINE=2010255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA      Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA      White O., Hickey E.K., Peterson J.D., Utterback T.R., Barry K.J.,
RA      Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA      McClarty G., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA      McClardy G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT      "Genome sequences of Chlamydia trachomatis Moyn and Chlamydia
RT      pneumoniae AR39.";
RL      Nucleic Acids Res. 28:1397-1406(2000).
RN      [4]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=J138;
RC      MEDLINE=2030349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA      Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA      Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT      "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT      from Japan and CWL029 from USA.";
RL      Nucleic Acids Res. 28:2311-2314(2000).
RN      [5]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=TW-183;
RA      Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA      Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT      "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT      other Chlamydia strains based on whole genome sequence analysis.";
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC      (Potential).
CC      -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC      -1- CAUTION: Ref.4 and Ref.5 sequences differ from that shown due to a
CC      frameshift in position 673.
CC
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      EMBL; AJ133035; CAB37083.1; -; Genomic DNA.
CC      EMBL; AE001586; AAD18172.1; -; Genomic DNA.
CC      EMBL; AE002235; AAF38561.1; -; Genomic DNA.
CC      EMBL; BA000008; BAA98223.1; ALT_FRAME; Genomic DNA.
CC      EMBL; AE017157; AAP97948.1; ALT_FRAME; Genomic DNA.
CC      EMBL; AE017157; AAP97949.1; ALT_FRAME; Genomic DNA.
```

DR PIR: E72130.
DR PHCI-2DPAGE; Q9Z3A1; --
DR TIGR; CP0761; --
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR011427; ChlamPMP.M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR Pfam; PF07548; ChlamPMP.M; 1.
DR TIGRfam; TIGR01376; POMP_repeat; 3.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 841 Probable outer membrane protein pmp2.
FT CONFLICT 784 784 N -> S (in Ref. 4).
FT SEQUENCE 841 AA; 89601 MW; 906460D0678D24C CRC64;
SQ
Query Match 17.8%; Score 939.5; DB 1; Length 841;
Best Local Similarity 27.0%; Pred.No.6e-47;
Matches 289; Conservative 134; Mismatches 358; Indels 291; Gaps 30;
QY 1 MOTSFHKPLSMILAYSCCSLNGGGYAAIMVPOGIYDGETLTVSPYTVIGDPGCTTVF 60
DB 1 MKIPLFLLSLIVPTLSMNLGAATTELSASNS-FDGTSTTSFSSKTSSATDGTNV 59
QY 61 SAGELTLKLDNSIAALPLSCFNG--LLGSFTVLGRHSLTPENI-RTSTNGAALSNSAA 117
DB 60 FKDSVVIENVPKTGETQSTSCFKNDAAGDLNPLGGGFTFSNIIDATTASGAIGSEAA 119
QY 118 DGLFTIEGFKELSFSCNSLLAVLPAATTNKGSTPTTTSTPSNGTIYSKTDILLNNEK 177
DB 120 NKTVLGFSALSF-----LKPASTVTNGLGAINVKGNLSLLDNDK 161
QY 178 FSYSLNVLGSGDAID-AKSLTVQGISKLCVFOENTAQADGGACQVTSFSAWANEAPTA 236
DB 162 VLIQDNFSTGCGAINCAGSLKI-----ANNKLS 191
QY 237 FVANAGVGGGTAQVQDGGQVSSSTSTEDPVVFSRNRTAVEFDGNVARVGGIYSYGN 296
DB 192 FIGNSSSTRGAIHT-----KN 208
QY 297 VAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGGAI FCKNGAAGSNNSGV 356
DB 209 LTLSSGGETLFGNTA-----PT----- 226
QY 357 SPDEGVWFSSNVAAGKGAIVAKKLSVANGCPVQFLRIANDGGAIVLGESGELSLSA 416
DB 227 -----AAGK-----GGAIAIADSGTSLISG 246
QY 417 DYGDIIIPDGNLKTAKENADVNGVTYSSQAI SMGSGGKLTIRAKAGHOILPNDPIEMA 476
DB 247 DSGDIIIEG-----NTIGATG-TVSHSAIDLGTSAKITALLRAAQGHYIYFDPIIV- 296
QY 477 NGNNQPAQSSKLLKIN-----DGEVYTGDI VPA-----NGSSTLYQNVITIQ 518
DB 297 TGTSVADA---LNINSPDTGDNKEYTGTIVFSGEKLTEAEAKDEKRTSKLLQNVAFKN 353
QY 519 GRIVLREKAKLSVNSLSQTGGS-LYMBAGSTWDPVTPPQQPPAANQLITLNLHLSLS 577
DB 354 GTVVLKGDVLSANGFSQDANSKLIMDLGTSL-----VANTESIELTNLEINID 402
QY 578 SLIANNVATPPTNPAPQDSHPAVIGSTTAGSVTISGPIPEDLDDTAYDR-----Y 629
DB 403 SLRNGKKI-----KLSAATAQKDIRIDRPVVLATSDSFYQNGFLNEDHSY 448
QY 630 DMLGSNQKINVLKQLGKTPANAPSDLTLGNEMPKYGYGQSWKLDWDPNTANNPPTLK 689
DB 449 D-----GILEDAQKDIVISADRSIDAVQSP-YGYQKWTINSTDDK-----K 492
QY 690 AT--WTKTGNPGERVASLVPNSLWSIILDIRSAHSAIAQASVDGRSYCRGLWYGSVNF 747
DB 493 ATYSNAKQSNPTAEQAEPVLPNLWGSFIDVRSFQNFIELGTGAPYKRFWVAGISNV 552
QY 748 FYHORDALGQGYRYISGYSIGANSYF-GSSMFLGAPTEVFGRSKDVVVCNSNHACIGS 806

DB 553 LHRSGRENQRKFRHVSGGAVVGASTRMPGDDTSLGFAQLFARDKDYFMNTNFAYKYS 612
QY 807 VYLSTQOALCG--SVLFGDAFIR-----ASYCFGHQMHTSY----- 841
DB 613 LRLQHDASLYSVVSLILGEGGLREILLPYVSKTLPSCSPYQGLSYGHTDHRMKTESLPPPP 672
QY 842 -TFAEESDVRWDDNCLAGEIGAGLPVITPISKLYLNEIRPPVQAFSADHESFTEEGDQ 900
DB 673 PTLSTD-HTSNGYVWAGELGTRVAVENTSGRFPFQEVTPFVKVQAVARQDSFVELGAI 731
QY 901 ARAFKSGHLLNLSVPVGVKFCRCSSTHPNKYSFMAAYICDAVRTISGTTETLLSHOETWT 960
DB 732 SRDPSDLSHLYNLAIPLGIKLEKRFPA--EQYHVVMVAMYPDVCRSNPKCTTTLLSNQGSWK 789
QY 961 TDAFHARHGVVVRGSMVASLTSTNIEVYGHGVEYRDASRGYGLGAGSRVRF 1012
DB 790 TKGSLNARQAGIVQASGRFSRLGAAALFNGFGEWRGSRSTNVNAGSKIKP 841
RESULT 24
P71133_CHLAB
ID P71133_CHLAB PRELIMINARY; PRT; 846 AA.
AC P71133;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE POMP91B precursor.
OS Chlamydia abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S26/3;
RX MEDLINE=98187897; PubMed=9529048;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "Molecular cloning and characterization of the genes coding for the
RT highly immunogenic cluster of 90-kilodalton envelope proteins from the
RT Chlamydia psittaci subtype that causes abortion in sheep.";
RL Infect. Immun. 66:1317-1324 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S26/3;
RX MEDLINE=96406378; PubMed=8810511; DOI=10.1016/0378-1097(96)00281-9;
RA Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RT "Identification of a multigene family coding for the 90 kDa proteins
RT of the ovine abortion subtype of Chlamydia psittaci.";
RL FEMS Microbiol. Lett. 142:277-281 (1996).
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto_transpbeta.
DR InterPro; IPR011427; ChlamPMP.M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF07548; ChlamPMP.M; 1.
DR Pfam; PF02415; Chlam_PMP; 3.
DR TIGRfam; TIGR01376; POMP_repeat; 3.
KW Signal.
FT SIGNAL 1 16 Potential.
FT CHAIN 17 846 POMP91B.
FT SEQUENCE 846 AA; 90834 MW; 4CDC31DC03C2964E CRC64;
SQ
Query Match 17.0%; Score 897; DB 2; Length 846;
Best Local Similarity 27.0%; Pred.No.2e-44;
Matches 278; Conservative 133; Mismatches 353; Indels 264; Gaps 30;
QY 40 ETITVSPFY--TWIGD-----PSGTTVPSAGELTLKLDNSIAALPLSCFNGLLGSFT 90
DB 28 ETITSSDSNGVNTSDEFEVKETTSAGIYTCGNCVLSIYAGKD-SPLNKSCFSETTENS 86
QY 91 VLGRGHSLTPENIRT-STNGAALSNSAADGLFTIEGFKELSFNCNSLLAVLPAATTNKG 149
DB 87 FIGNGYTLCFDNTTQSSHPGAISVSTNKTLDISGSLFSCAYC-----CPGTTGYG 140


```

Db      358 KSEFQKQKLLFWHFGT-----PLQTNNL-SIKNLHLDSKIAA-TAV----- 399
Qy      591 NPQAQDHPAVTGSTTAGSVTISGPIFFEDLDATYDRYDVMGSGNQKINVLKQLGTGKPP 650
Db      400 -----EIAATAAIAIEICGPMVWH-VDEIFPYNQALADSLSPFCLHVR----- 443
Qy      651 ANAP--SDTLGNEMP-----KYVGQSGWKLAWDP-----NTANNPGYTLKATW 692
Db      444 --APHLDNITV-DDVPLIPITTMETIRGYQGWTSWEEHEHMFQNVSTQPNKMGSLVW 500
Qy      693 TTKGYNP-----GPERVASLVPNSLWGSILDIISAHSIAIQASVDGRSYCKRGLWVGVSYPNF 748
Db      501 NPSGYIPFVGCGEFTTSLVPSNLWNLFLDTFAQAQAI--TNAQSPGNGIWISSLTNSF 558
Qy      749 YHRRDALGQYRYISGYSGLAN-SYFGSSMFLGAPTEVFGRSKDYVVCRRSHHACIGSV 807
Db      559 RKGSTENHGFRRKSSGYVAGGKFTQLDDIFSVGICQLFGRSKDFGSAKSKAPFSGSL 618
Qy      808 YLSTQOAL--CGSYLFG-----DAPIRASVGFQGNHMKTSYTPAE 845
Db      619 YAHHSRYLLPITRFLAGTSGRSQRFSLRIPKDPPIINFDAI--SYSGRNHMKVYADHS 676
Qy      846 ESDVRDNNCLAGEICAGLPVITPCKLYLNLRLPFWQAEFSYADHESFTEGDAQARAFK 905
Db      677 QTTSWNTGYSAQIGSSLPCLDVSHTFPQVSPFVKLHWIYAHQVQFQEGIKERSEN 736
Qy      906 SGHLLMSLPVGVKPRCSSTHPNKYSFMAAYICDAYRTISGTFTTLLS-HOBTWTTDAP 964
Db      737 NSNLKNSLPGLKI--QGSLHLHSYELTGMVYADLYRCNPESVTSLSGGLLPWTTTAA 795
Qy      965 HLAHGVVVRGSMYASLTSNIEVYGHGRYEDRSDRGYGLSAGSRVFP 1012
Db      796 NLGKQALLQSGNLSLTSHINIFAQGTVEFRSSYSYAMDAGSRVHP 843

RESULT 26
PMPI_CHLMU      STANDARD;      PRT;      867 AA.
ID      PMPI_CHLMU
AC      Q9PL41;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-MAY-2005 (Rel. 47, Last annotation update)
DE      Probable outer membrane protein pmpI precursor (Polymorphic membrane
      protein I).
GN      Name=pmpI; OrderedLocusNames=TC0267;
OS      Chlamydia muridarum.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=83560;
RN      [1]
RC      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX      STRAIN=Mopn / Nigg;
RX      MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA      Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA      White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA      Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA      Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA      McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT      "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
      pneumoniae AR39."
RL      Nucleic Acids Res. 28:1397-1406(2000).
CC      -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
      (Potential).
CC      -1- SIMILARITY: Belongs to the pmp outer membrane protein family.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; AE002294; AAF39136.1; -, Genomic_DNA.
DR      PIR; F81721; F81721.

```

```

DR      TIGR; TC0267; -.
DR      InterPro; IPR005546; Auto transpbeta.
DR      InterPro; IPR011427; ChlamPMP M.
DR      InterPro; IPR003368; Chlamydia_PMP.
DR      Pfam; PF03797; Autotransporter; 1.
DR      Pfam; PF02415; Chlam PMP; 6.
DR      Pfam; PF07548; ChlamPMP M; 1.
DR      TIGRPFAMs; TIGR01376; POMP repeat; 6.
KW      Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT      SIGNAL      29      Potential.
FT      CHAIN      30      867      Probable outer membrane protein pmpI.
SQ      SEQUENCE      867 AA; 95018 MW; 557994185A9E5652 CRC64;

Query Match      15.0%; Score 789; DB 1; Length 867;
Best Local Similarity      26.1%; Pred. No. 5.2e-38;
Matches 268; Conservative 148; Mismatches 411; Indels 200; Gaps 31;

Qy      18 CCSLNGGGAARIMVPOGIYDEGTUTVAFPYTVIGDPSGTTFVSAGELTKLNDLSIAL 77
Db      9 CC-----LCATILSPTAILFGQD-ALDKSALITKNPNSIVCTFLEDCTMENFSPALLSH 61

Qy      78 -----PLSCFNGLLGSFTVLGRGHSILTPENIRTSNGAALSNSAADGLFTIEGFKELSFS 132
Db      62 ARQDDPLVIIGNT-----HWFVSNLHPSTNEERFLKEKGD--LSIQDFRFLST 109

Qy      133 NCNSLLAVLPAATTNKGSGTPTTTPNGNTIYSKT-DLLLLNNEKFSFYNLVSDGGA 191
Db      110 DCS-----STEDSPS--ILYHKGQLFLRNGNWSFYRNHSESGGA 150

Qy      192 IDAKSLTVQGISKLCVFOENTAOAGGACQVVTSPFSAMANEAPIAFVANAVGVRGGIAA 251
Db      151 LSTDALFLQHNLYFTNFENSSAKNGAIOAQT----- 183

Qy      252 VDDGQGVSSSTSTEDPVVPSRN-TAVEFDGNVAVGGIYSGYGNVAFNLNKGKTLPLNN 310
Db      184 -----LSLRNVSLSFSNRNRLNGAICCCQLNLCGNVNLPIFF-- 223

Qy      311 VASPVYIAAKQPTSGASNTSNYDGGAIKCKNGAQAAGSNNSGVSVDGEGVVFSSNV 370
Db      224 -----TNSALNGGAIKCCINEQLSEKCLSLAYNQ--TLFSGNS 262

Qy      371 AAKGGAIYAKKLSVANCGPVQLRNANDGGAIIYGESGELSLSDYDIIIPDGNLKR 430
Db      263 AKEKGAIIYKHWLRHNGPVSFVNSAKLGAIAIQSGGSLIIAGGSGVLPQNNSCHP 322

Qy      431 AKENAAVNGVTVSSQAISMGGKITTIRAKAGHOILFNDPIEMANGNNQPAQSKLLK 490
Db      323 SDQGTV-----RNAIYLEKNALLSSLEARHG-DILFFDPPIVQEVVSEFSTTSALT 373

Qy      491 INDGEGYTGDIIVP-----ANGSSTLYQNVITBQGRIVLREKAKLSVNSISQTG 538
Db      374 LRIQTWTRAVIFSSNLSKEEKEATEANLSKIQPIELQSGCLVKDRVILSAPLSQAP 433

Qy      539 GS-LYNEAGSTWDFVTVPQPPQPPAAANQLITISNLHLSLSSLLANNNAVNTPTNPQAQDS 597
Db      434 QALLVNDVGT-----LTTSDKLTLTSLPLHSI----- 463

Qy      598 HPAVIGSTTAGSVTISGP-----IFFEILD-DTAYDRDVLGSKNQ-INVLKQLGTGK 649
Db      464 -----DTENSVSISQPTLSIQKIFLSNSENHYENVLLSKQKQDIPLLSLPLGLPH 516

Qy      650 PANAPSLDITGNEMPKYGVGSKWKLAWDPNTANNPGYTLKATWTKGYNPGPERSVLVP 709
Db      517 P-----DLPDGNLSHFQYQGDWNSWQTSQDQRE---TLVANVTANSIYIPHERQSALVA 568

Qy      710 NSLWGSILDIRSAHSAIQASVDGRSYCRGLWVGVSYPNFY-HDRDALG-QGYRYISGGYS 767
Db      569 NTLWNTYSDMQAVQSMINTTAQCGAVLFGTWGSAVSNLFFSHGNSGKSTDNWKRSLGYL 628

Qy      768 LGANSY-FGSSMFLGAPTEVFGRSKDYVVCRRSHHACIGSVYSTQALCGSYLFGDAFI 826
Db      629 FGISTHSLDDHSFCLAAQQLFGKSSDFVTSADTTSYIAAI-----QTQIATSLIKISA-- 682

```



```

Qy 827 RASVGFQGHMKTYS-TFAESDVRWNNCLAGEIGAGLPVITPSKLYLNELRPFVQAE 885
Db 693 QACYNESHELTKTYSKSGFGAWSHVAVSGEIGASIPVNSGGLF-SFSIFSJKLQ 741
Qy 886 FSYADHESFTBEGDQARAFKSGHLNLNLSVPVGVKFDRCSSSTHPNKYSFMAAYICDAVRTI 945
Db 742 GFSGQKQDFEBSRGEARAFADSSFTNISLPVGAIEPKKSQKTRNYHFLGAYIQDLKCV 801
Qy 946 SGTETILLSHGETWTTDFLPHARHGVVVRGSMYASLTNIEVYGHGRYEYDASRGYGLS 1005
Db 802 ESGPVTLKNSVTWDAPMANLDSRAMFRLTNORAL-HRFQTLNMSYMLRGQSYSYSLD 860
Qy 1006 AGSRVRF 1012
Db 861 LQTYRPF 867

RESULT 27
Q8VU49 CHLPS
AC Q8VU49; CHLPS PRELIMINARY; PRT; 602 AA.
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative polymorphic membrane protein (fragment).
OS Chlamydia psittaci (Chlamydiales; Chlamydiaceae; Chlamydia).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LLG;
RA Laroucau K., Souriau A., Rodolakis A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243418; AAL36962.1; -; Genomic DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR006315; Auto transporter.
DR InterPro; IPR005546; Auto transporter.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR TIGRFAMs; TIGR01414; Autotrans_bar1; 1.
FT NON TER
SQ SEQUENCE 602 AA; 65561 MW; CA486CFACEC131E2 CRC64;

Query Match 15.0%; Score 787.5; DB 2; Length 602;
Best Local Similarity 33.0%; Pred. No. 3.8e-38;
Matches 210; Conservative 99; Mismatches 250; Indels 77; Gaps 21;

Qy 419 GDILPDGNLKTAKENADVNGVTSSQAIISWGGKITTILRAKAGHQILFNDIEMANG 478
Db 2 GDITPDGNKIITTSRSSS-----TVKRNISISLGGKFTKLNAKEGFIFFYDPIANTGD 56
Qy 479 NNQPAQSGLKLIKNDGEG-----YTGDIVFA-----NGSSTLYQNVTIEQGR 521
Db 57 TNYE-----TELNAEGGSTTYTGKIVFSEKLSDEKVDNLKSYFTQPLKIGAGSL 110
Qy 522 VLREKAKLSVNSLSQSGS-LYMEAGSTWDFVTFPQPPQPPAANQLITLSNLHLSLSLL 580
Db 111 VLKDGVTLEAKKVQGTGISTVMDLGT-----LQTSSTGRTITLNLNDINVAS-L 160
Qy 581 ANNAVTPNPPTPAQDSHPAVIGSTAGSVTISGIPPELDLDTAYDRYDHLGNSKINV 640
Db 161 GGGGVADPFAKVAQAS-----GKT-----VTINA-VNLVDVDGNAYE-YPLATSQPPTA 209
Qy 641 LKLQLGKPPANAPSDTLGNEMP--KYGYGSGKLAWDPNANNGPYTLKATWTKTGYN 698
Db 210 IIAKAGSGTITTFD-NLKNYTPHYGQNTVTVKLGTSQA-BETATLTWEQTDYS 267
Qy 699 GGPBRVASLVNSLWGSILDIRSAHSAIQASVQGRSYCRGLWVSGVGNFFYHDDRDLGQG 758
Db 268 FNPBRQGLVPNTLWGSFSDIRAIQNLIDISVNGADYRRGFWVSGLGNFLHKSGNSTKR 327
Qy 759 YRISGGYSLGANSYFGS-SNFGLAFTVFGRSKDYVYVCRSNHHACIGSVYL-----ST 811

```

```

Db 328 FRHSAGYALGVTAQTSTEDVFSAAFQCFQKQKDYFVKNSSNIYAGSIYQHISYNA 387
Qy 812 QQALCGSYLFGDAFI-----RASVGFQGHMKTYS-----PABESDVRWNNCLAGE 859
Db 388 WONLQSTIGAEAPLVLAQLTYCHASNNKNTNTVVPKNVTLSEIKGDMGNDCFQGE 447
Qy 860 IGAGLPVITPSKLYLNELRPFVQAEFSYADHESFTBEGDQARAFKSGHLNLNLSVPVGV 918
Db 448 FGAMAPIE-TPSSFLFDYRSPFLQLQLVHAHQDDFKENNSDQGRYPFESSNLTNLSMPTGI 506
Qy 919 KFDRCSSSTHPNKYSFMAAYICDAVRTISGTETILL--SHQETWTTDFLPHARHGVVVRGS 976
Db 507 KPERFANDVASVHLTAAYAPDIVRNPDCATSLVSPFSAVWVTKANNLARSAFMLQAG 566
Qy 977 MYASLTNIEVYGHGRYEYDASRGYGLSAGSRVRF 1012
Db 567 NYLALSHNMLFSQFGFEIRGSSRTYNDLGSKIQF 602

RESULT 28
Q823X3 CHLCV
ID Q823X3 CHLCV PRELIMINARY; PRT; 843 AA.
AC Q823X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymorphic outer membrane protein G family protein/autotransporter, putative.
DE OrderedLocusNames=CCA00280;
GN OrderedLocusNames=CCA00280;
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GPIC;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T., Heidelberg J.F., Holtzapfle E.K., Khouri H.M., Federova N.B., Carty H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J., White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M., Fraser C.M.;
RA "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae."
RT Nucleic Acids Res. 31:2134-2147(2003).
RL EMBL; AE016995; AAP05031.1; -; Genomic DNA.
DR TIGR; CCA00280; -;
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transporter.
DR InterPro; IPR011427; ChlamPMP_M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; IPR006626; PbH1.
DR Pfam; PF03797; Auto transporter; 1.
DR Pfam; PF07548; ChlamPMP_M; 1.
DR Pfam; PF02415; Chlam_PMP; 4.
DR SMART; SM00710; PbH1; 2.
DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
KW Complete proteome.
SQ SEQUENCE 843 AA; 90476 MW; 0BB240F4687AFC7A CRC64;

Query Match 14.9%; Score 785.5; DB 2; Length 843;
Best Local Similarity 26.1%; Pred. No. 8.1e-38;
Matches 270; Conservative 131; Mismatches 351; Indels 281; Gaps 37;

Qy 37 YDGETLTVSPYTVIGDPSTGVFVPSAGELTKMLNLSIALPLSCFNLGSGFTVLRGH 96
Db 35 FDSGTAGQRTSKQSTDAGGTTTTLGDVTIIVTKTSPA-NTSCFNSGNTLFTGANH 93
Qy 97 SLTFENITSTNGAALNSAADGLFTTEGPKELSFNSCNLSLLAVLPATTNKGSQTPTTT 156
Db 94 SLIFEDIISTAQGAALIS-----ANT----- 113

```



```

Db 530 LWNFLPLDTRFQQAIEK--HAVSSGNGIWISSMTNNSPLQGGSTNNHNGFRHKSSGYTAGGK 587
Qy 772 -SYFGSSNPFGLAFTVFRGSKDYVVCRRNHACIGSVYLSQQALCG--SYLFGDAFTR- 827
Db 588 IQTLQDDIFSVFSQPKRSKDFGSATSKDTFLSGSIYAQHSRRLLPIMRFLAGTSTYRP 647
Qy 828 -----ASYGFGNOHMKTSYTFABESDVRMNNCLAGEICAGLPVITTP 870
Db 648 RLLISIPKXLPINFVLVSYSYSDSNHMKVKNSTQTGRGWNTRYGSAQIGSSIPFLVDV 707
Qy 871 SKLYLNELRPFVQAFESYADHESFTBEGDQARAFKSGHLLNLSVPVGVKFPDRCSSTHPNK 930
Db 708 SHTFPQYVSPFKLHWIAHQVQFOEQGKRRSPNNSLNKLSLPLGLKI-QGQTLRLS 766
Qy 931 YSFMAAYICDAYRTISGTETILLS-HQETWTTDAPHLARHGVRVGRGSMYASLTNIEVYG 989
Db 767 YEFTGMWITDLYRCNPEGVSTLSIGLLPWTITGTNLSKQAALLQGSNLSLTSHINIFA 826
Qy 990 HGRVEYRDASRGYGLSAGSRVRP 1012
Db 827 QGTVEFRSSYSYNNMDFGSRVHP 849

RESULT 30
Q4VWS6_CHLTR
ID Q4VWS6_CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWS6
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TW-3, TW-448, TW-5, and Apache-2;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299430; AAQ74446.1; -; Genomic DNA.
DR EMBL; AY299442; AAQ74458.1; -; Genomic DNA.
DR EMBL; AY299448; AAQ74444.1; -; Genomic DNA.
DR EMBL; AY299449; AAQ74445.1; -; Genomic DNA.
DR InterPro; IPR005546; Auto.transp.beta.
DR InterPro; IPR011427; ChlamPMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; Chlam_PMP; 6.
DR Pfam; PF07548; ChlamPMP M; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
SQ SEQUENCE 878 AA; 95489 MW; 9E7611D9B0B636F7 CRC64;

```

```

Query Match 14.9%; Score 784.5; DB 2; Length 878;
Best Local Similarity 26.5%; Pred. NO. 9.8e-38;
Matches 267; Conservative 146; Mismatches 392; Indels 203; Gaps 33;

Qy 53 DPSTGVTFPSAGELTKNLDNSIAALPLSCFNGLL-----GSPVLVGRGSLTFE 101
Db 26 DPLGETA-----LTKPNHVVCVFEDCTWESLFPALCAHASQDDPLVVLGNSYCWFS 80

Qy 102 NIRTSTGAALNSAAGDLFTIEGKELSPNCNSLIAVLPAATNKGSPPTTTPSN 161
Db 81 KLHITDPKEALFKEGD--LSIQNFRFLSPFDCSS-----KESSPS- 119

Qy 162 GTIYSKT-DLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFQENTQAQDGAC 220
Db 120 -IIHKQKQLSLRNNGNSFCRNHABGGGAISADAFSLQHNLYLFTAFPESSKNGGAI 178

Qy 221 QVWTSFSAWANEAPIAFVANYAVAGVGGGIAAQVQDQGGQSVSSSTSTEDPVVFSRNTAVEF 280
Db 179 QAQT-FSLSRNVSPISFARNADLNGGAICC----- 208

```

```

Qy 281 DGNVARVGGIYSGYCNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYDGGAI 340
Db 209 -----SNLICSGNV-----NPLPF-----TCNSATNGCAI 233
Qy 341 FCXGQAQAGSNNSGSVFDGEGVVFPSSNVAAGKGAIAKAKLSVANCGPVQFLRNLIAND 400
Db 234 CCI--SDLNTSEKGSLSLACNQETLFPASNAKEKGAIAKAMVLRVNGPVSPFINNSAKI 291
Qy 401 GGAIYLGESGELSADYGDIIIFDGNLKTAKENAADVGVTVSSQAISMGSGGKITTIR 460
Db 292 GGAIYAGGSLSLAGEGSLVLFQNNQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343
Qy 461 AKAGHQILFNDPI-EMANGNNQPAQSSKLLKINDGEGYT-----DIVPA----- 504
Db 344 ARNG-DILFDPPIVQESSKESPLSSIQASVTSPTPATASPLVIQTSANRSVIFSGERL 402
Qy 505 -----NGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTTGS-LYMEAGSTWDFVTPQ 555
Db 403 SEBEKTPDNLTSQLPIELKSGRLVKDRAVLASPLSDOPQALLIMEAGTS----- 455
Qy 556 PPOQPPAANQIITLSNLHLSLSLLANNAVTPNPTPPAQDSHPAVIGSTTAGSVTISGP 615
Db 456 -----LKTSSDLKATLSIPLHSL-----DTEKSVTIHAP 485
Qy 616 -----IPFEDL-DDTAYDRYDMLGSNQKINVLKQLGTQPPANAPSDLTIGNEMPKYGY 668
Db 486 NLSIQKIFLNSGDNENFYENVLLSKEQN-----NIPLLTLPKEQSHLHLPDGNLSSHFGY 541
Qy 669 QGSMKLAWDPTANNNGPYTLKATWTKGYNPGPERVASLVPNSLWGSILDIRSAHSAIOA 728
Db 542 QGDWTFSW--KDSDEG-HSLIANWPKYVPHPERQSTLVANTLWNTYSMDQAVQSMINT 598
Qy 729 SVDRGSRGVLWVGVSNNFFY-HDRDALG-QGYRIISGGYSLGANSY-FGSSMFGLAFT 785
Db 599 TAHGAYLFGTWGSAVSNLFAHDSGKPIDNWHHRSGLYLGISTHSLDDHSECLAAQ 658
Qy 786 VFGSKDYVVCRRNHACIGSVYLSQQALCGSYLFGDAFIRASYFGNGHMTSY-TPA 844
Db 659 LLGKSSDSFITSTE-----TTSVIATVQAQLATSLM-KISAQACYNESIHELKTKYRSFS 712
Qy 845 EESDVRNDCNCLAGEICAGLPVITTPSKLYLNELRPFVQAFESYADHESFTBEGDQARAP 904
Db 713 KEGFGSMHVSVAVSGEVCALPIVNSGSLF-SSFSIFSKLQGFSGTQDGFESSGHIRSP 771
Qy 905 KSGHLLNLSVPVGVKFPDRCSSTHPNKYSFMAAYICDAYRTISGTETTLTLLSHQETWTTDAF 964
Db 772 SASSFRNISLPIGITPEKKSQKTRNYTFYLGAYIQDLKRDVESGPPVLLKNAVSMADPMA 831
Qy 965 HLAHGVVVRGSMYASLTNIEVYGHGRYERDRASRGYGLSAGSRVRP 1012
Db 832 NLASRAYMFLTNQAL-HRLQTLNLSVYVLRGQSHSYSLDLGATHRP 878

```

```

RESULT 31
Q4VWR2_CHLTR
ID Q4VWR2_CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR2
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bour;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299443; AAQ74459.1; -; Genomic DNA.

```

SQ		SEQUENCE	878 AA; 95629 MW; CA385BFC8BAC1C17 CRC64;
Query Match		14.8%; Score 781.5; DB 2; Length 878;	
Best Local Similarity		26.6%; Pred. No. 1.5e-37;	
Matches		269; Conservative 146; Mismatches 388; Indels 207; Gaps 35;	
QY	53	DPSGTTVFSAGELTILKNDLSIAALPLSCFGNLL-----GSFTVLGRGHSLTPE	101
DB	26	DPLGETA-----LLTKNPNHVVCTTFEDCTMESLPALCAHASQDDPLYVLGNSCWVFS	80
QY	102	NIRTSNGAALSNSAADGLFTIEGPKELSFNSCNSLLAVLPAATTNKGSTPTTTSTPSN	161
DB	81	KLHITDPKEALPKEGD--LSIQNFRFLSFTDCSS-----KSSSPS-	119
QY	162	GTIYSKT-DLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOENTAOADGGAC	220
DB	120	-IIHQKNGQLFLRNNGSMFSCRHAEGSGAISADAFSLQHNYLFTAPEENSKNGGAI	178
QY	221	QVVTSPSAMANEAPAFVANVAGVGGGIAAVDQGGQGVSSSTSTEDPVVPSRNTAVEF	280
DB	179	QAQT-FSLSRNVSPISFARNRADLNGAICCC-----	208
QY	281	DGNVARVGGGIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGOASNTSNNGDGGAI	340
DB	209	-----SNLICSGNV-----NPLFF-----TGSATNGGAI	233
QY	341	FCKNGAAGSNNSGVSFDGEGVFPSSNVAAGKGAIIYAKKLSVANCGPVOFLRNAND	400
DB	234	CCI--SDLNTSEKGSLSLACNQETLFPASNSAKEKGAIIYAKHMLRYNGPVSFINNSAKI	291
QY	401	GGAIYLGESGELSADYGDIIFDGNLKTAKENADVNGVTVSSQAI5MGSGGKITTLR	460
DB	292	GGAIATQSGGSLILAGEGSLVFQNNRSQRTSDQGLVR-NAIYLEKDAI-----LSSUE	343
QY	461	AKAGHOILFNDPI-EMANGNNOPOASSKLLKINDGEGYTG-----DIVFA-----	504
DB	344	ARNG-DILFPDPIVQSSSKESPLPSSLOASVTSPTPATASPLVIQTSANRSVIFSSERL	402
QY	505	-----NGSSTLYQNVTIEQRIYVLRKAKLSVNSLSQTGGS-LYMBAGSTWDPVTPQ	555
DB	403	SEEEKTPDNLTSQLOQPIELKSGRLVKDRAVL5APLSQDPOALLIMEAGTS-----	455
QY	556	PPQOPPAANOLITLSNLHL5LSLLANNVNTPTNPPAQDSDHPAVIGSTTAGSVTISGP	615
DB	456	-----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP	485
QY	616	-----IFPEDL-DDTAYDRYDMLGSNQKINVLKLOLGTKPPANAPSDLT--GNEMPKY	666
DB	486	NLSIQIKFPLNSGDNFYENVELLSKEQN-NIPLLT-----SKEQSHLHLPDGNLSSH	539
QY	667	GYQGSWKLAWDPNTANNGPYTLKATWTCTGYNPGERVASLVPNSLWGSITLDRSAHSAI	726
DB	540	GYQGDWTFW--KDSDEG-HSLIANWTPKNYVPHPERQSLTAVLNTWYSDMQAVQSMI	596
QY	727	QASVGRSICRGLVWGSVSNFFY-HRDALG-QGYRIYISGGYSLGANSY-FCSSMGLAF	783
DB	597	NTTAHGAYLFGTWGSAVNSLFLYANDSGKPIDNWHRSISGLYLFGISLTHSLDDHSPCLAA	656
QY	784	TEVPGRSKDVVCRSNHHACIGSVLSTQALCGSLFGDAPFIRASVFGNGHMTKSY-T	842
DB	657	GQLLGSSDSFITS-----TTSYIATVOAQLATSML-KISAQACYNESIHELKTKYRS	710
QY	843	FABESVRWNNNCIAGEIGAGLPIVTPPSKLYLNELPFPVQABFSYADHSFTEEGDQAR	902
DB	711	FSKEGFGSWHVA5GEVCASIPIVNSGGLF-SSPSIFSKLQFGSQDGFESSSGEIR	769
QY	903	AFKSGHLNLT5VPVGVKFDRCSSTHPNKYSFMAIYICDAVRTISGTETTLLSHQETWTD	962
DB	770	SFSASSPRNLSLPIGTIFTEKKSQKTRNYIYFLGAYIQDLKRDVKSGSPVTLTKNAVSAL	829
QY	963	APHLARHGVVVRGSMYASLTSNI5EVYGHGRYEYEDASRGYGLSAGSRVF	1012
DB	830	MANLDSRAYMFLRLNQAL-HRLQTLNLVSVLGRQSHSYSLDLGTTTRF	878

```

Db      540 GYQGDWTFW--KDSDEG-HSLIANWTPKNYPHPERQSTLVANTLWNTYSMDQVQSMI 596
Qy      727 QASVDGRSYCRGLWVSGVSNPFY-HRDALG-QGYRYISGGYSLGANSY-FGSSMFLGAF 783
Db      597 NNTAHGGAYLFGTWSGAVSNLFYAHDSGKPIDNWHHRSLGSLYLGISTHSLDDHSPCLAA 656
Qy      784 TEVGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAFIRASYFGNQHKTSY-T 842
Db      657 GQLGKSDSPSTSTE-----TTSYIATVQALATSLM-KISAQACYNESIHELKTKYRS 710
Qy      843 FAEEDVDWNNCLAGEICAGLPIVITPSKLYNELRPFVQAEFSYADHESFTBEGDQAR 902
Db      711 FSKEGFGSHSVAVSGEVCASPIVNSGSLF-SFSPIFSKLQSGFTQDGFESSGEIR 769
Qy      903 APKSGHLLNLVPGVGVKFCRSTTHPKNSYMAAYICDAYRTISGTETTLSSHQETWTD 962
Db      770 SFSASSFNISLPIGITPEKKSQKTRNYVFLGAYIQDLKRDVKSPPVTLKNAVSWDAL 829
Qy      963 AFHLARHGVVVRGSMYASITSNIEVYHGRIYRDSRGGYGLSAGSRVRF 1012
Db      830 MANLDSRAYMFLTNQORAL-HRLQTLNLVSYVLRGQSHSYSLDLGTTYRF 878

RESULT 33
Q4VWR8 CHLTR
ID Q4VWR8 CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HAR-13;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299427; AAQ74443.1; -; Genomic DNA.
SQ SEQUENCE 878 AA; 95538 MW; 33F3F4B5F118AEECB CRC64;

Query Match 14.8%; Score 779.5; DB 2; Length 878;
Best Local Similarity 26.6%; Pred. No. 1.9e-37;
Matches 269; Conservative 148; Mismatches 385; Indels 209; Gaps 35;

Qy      53 DPGSTTVFSAGELTLKNDNSIAALPLSCFQNL-----GSTVLGRGHSLTPE 101
Db      26 DPLGETA-----LLTKNPNHVCTTFEDCTWESLFPALCAHASQDDPLYVLGNSYCWFS 80
Qy      102 NIRTSTNGAALNSAADLFTIEGPKELSFNCNSLLAVLPAATNKGSTPTTTSTPSN 161
Db      81 KLHITDPEALFKEGD--LSIQNFRFLSFTDCS-----KESSPS- 119
Qy      162 GTIYSKT-DLLLLNNEKFSYSLNLSVGGDAIDAKSLTVQGISKLVCVQENTAQDGGAC 220
Db      120 -IIHQKQLSLRNNGSMSPCRNHAHGSGGAISADAPSLQHNLYLPTAPEENSSKNGCGAI 178
Qy      221 QVTVSFSMANEAPAFVANVAVRGGGIAAVQDQOQGVSSSTEDTPVFSFRNTAVEP 280
Db      179 QAQT-FSLSRNVSPISFARNADLNGGAIC----- 208
Qy      281 DGNVARVGGIYSGYGNVAPLNNKTLFLNNVASPVYIAAKQPTSQAQNTSNNYDGGAI 340
Db      209 -----SNLCSGNV-----NPLFF-----TGNSATNGAI 233
Qy      341 FCKNGAAGSNNGSVSGVDFGEGVFPSSNVAGKGGAIYAKKLSVANGCPVQFLNLIAND 400
Db      234 CCI--SDLTSEKSGSLACQETLFASNSAKGGAIVAKHVLRYNGVPVFINNSAKI 291
Qy      401 GGAYILSGSGELSLSDYGDIIFDGNLKRATKENAADVNGVTVSSQALSMGSGGKITTLR 460

```

```

Db      292 GGATATGGSLSLAGEGSLVLPONNSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343
Qy      461 AKAGHQILFNDPI-EMANGNNQPAQSSKJLKINDGEGYTG-----DIVFA----- 504
Db      344 ARNG-DILFPDPIVQESSKESPLPSSLQASVTSPTATASPLVIQTSANRSVIFSSERL 402
Qy      505 -----NSSTUYQNWITIEQGRIVLRKAKLSVNSLSQGG-S-LYMEAGSTWDFVTPQ 555
Db      403 SEBEKTPDNLTSQLQPIELKSGSLVLDRAVLASPLSQDPPQALLIMEAGTS----- 455
Qy      556 PPQPPAANQLITLSNLHLSSLLANNVTPNTPPAQDHPAVIGSTTAGSVTISGP 615
Db      456 -----LKTSSDLKLTATLSIPLHSL-----DTEKSVTTHAP 485
Qy      616 -----IFFEDL-DTAYDRYDMLGSNQKINVLKQLGTKPPANAPSDLT--GNEMPKY 666
Db      486 NLSIQKIFLNSGDNFENVLLSKEQN-NIPLLT-----SKEQSHLHLPGDNLSSH 539
Qy      667 GYQGSWKLAMPNTANNPGPYTLKATWTKTGNPGPERSVASLVPSNLMSGSLDIRSAHAI 726
Db      540 GYQGDWTFW--KDSDEG-HSLIANWTPKNYPHPERQSTLVANTLWNTYSMDQVQSMI 596
Qy      727 QASVDGRSYCRGLWVSGVSNPFYHDDRDLGQ---GYRYISGGYSLGANSY-FGSSMFLA 782
Db      597 NNTAHGGAYLFGTWSGAVSNLFY-AHDSFGKPIDNWHHRSLGSLYLGISTHSLDDHSPCL 655
Qy      783 FTEVFGRSKDYVVCRSNHHACIGSVYLSLQALCGSYLFGDAPIRASYFGNQHKTSY- 841
Db      656 AQQLLGKSDSPITSTE-----TTSYIATVQALATSLM-KISAQACYNESIHELKTKYR 709
Qy      842 TPAESDVRRDNNCLAGEICAGLPIVITPSKLYNELRPFVQAEFSYADHESFTBEGDQA 901
Db      710 SFSKEGFGSHSVAVSGEVCALPIVNSGSLF-SFSPIFSKLQSGFTQDGFESSGEI 768
Qy      902 RAPKSGHLLNLVPGVGVKFCRSTTHPKNSYMAAYICDAYRTISGTETTLSSHQETWTT 961
Db      769 RSFSASSFNISLPIGITPEKKSQKTRNYVFLGAYIQDLKRDVESGFPVLLKNAVSWDA 828
Qy      962 DAPHLARHGVVVRGSMYASITSNIEVYHGRIYRDSRGGYGLSAGSRVRF 1012
Db      829 PMANLASRAYMFLTNQORAL-HRLQTLNLVSYVLRGQSHSYSLDLGATHRF 878

RESULT 34
Q4VWR8 CHLTR
ID Q4VWR8 CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Polymorphic membrane protein I.
GN Name=pmpI;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=UW-36, and UW-92;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY299437; AAQ74453.1; -; Genomic DNA.
DR EMBL; AY299435; AAQ74451.1; -; Genomic DNA.
SQ SEQUENCE 878 AA; 95568 MW; B0FD9FACB8A333AB CRC64;

Query Match 14.8%; Score 777.5; DB 2; Length 878;
Best Local Similarity 26.4%; Pred. No. 2.5e-37;
Matches 266; Conservative 145; Mismatches 394; Indels 203; Gaps 33;

Qy      53 DPGSTTVFSAGELTLKNDNSIAALPLSCFQNL-----GSTVLGRGHSLTPE 101
Db      26 DPLGETA-----LLTKNPNHVCTTFEDCTWESLFPALCAHASQDDPLYVLGNSYCWFS 80

```

QY	102	NIRTSNGAALSNSAADGLFTIEGPKELSPNSCNLLAVLPAATNKGSTPTTTSTPSN	161
DB	81	KLHITDPKEALPKKGD--LSIQNFRPLSFTDCSS-----KESSFS-	119
QY	162	GTIYSKT-DLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVQFQNTAQDGGAC	220
DB	120	-IIHQXGQLSLRNGSMSPFCRNHAEKSGGALSADAFSLQHNLYLTAPEENSSKNGGAI	178
QY	221	QVVTSPSAMANBAPAFVANVAGVGGGIAAVODGQQGVSSSTSTEDPVPVSPRTAVEP	280
DB	179	QAQT-FSLSRNVSPISFAHRADLNGCAIC-----	208
QY	281	DGNVARVGGGIYSYGNVAFPLNGKTLPLNNVASPVYIAAKQTSQASNTSNVYDGGAI	340
DB	209	-----SNLICSGNV-----NLPFF-----TGSATNGGAI	233
QY	341	PKNGAQAGSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANCGPVGQFLRNAND	400
DB	234	CCI--SDLNTSEKLSLACNQETLPASNAKEKGAIYAKHVLRYNGPVSFINNSAKI	291
QY	401	GGAIYLGESGELSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAIKMGSGKITTILR	460
DB	292	GGAIAIQSGSSILLAGESVLFQNNRSORTSDQGLVR-NAIYLEKDAI-----LSSLE	343
QY	461	AKAGHOILFNDPI-EMANGNQAQSSKLLKINDGEGYTG-----DIVPA-----	504
DB	344	ARNG-DILFPDPIVORSSSKESPLPSLQASVTSPATASPLVIOQTSANRVSIFSSERL	402
QY	505	-----NGSFLYQNVITEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDPVTPQ	555
DB	403	SEEEKTPDLNLSQLOPIELKSGRLVKORAVLSAFSLSQDPOALLIMEAGTS-----	455
QY	556	PPQOPPAANQLITLSNLHLSLLANNAVTPNPTNPPAQDHPAVIGSTTASGVTISGP	615
DB	456	-----LKTSSDLKLATLSLPLHSL-----DTEKSVTIHAP	485
QY	616	-----IPFEDL-DDTAYRDYMWLGSKNQKINVLKQLGTYPANAPSDLTILGNEMPKYGY	668
DB	486	NLSIQIFLNSGDNFYENVELLSKKQ-----NIFLLTLPKQSHVHLPDGNLSSHFG	541
QY	669	QSGKWLADPNTANNPYTLKATWTKGTGNPGERVASLVPNSLWGSILDIRSAHSAIQ	728
DB	542	QSDWTFSW--KSDSEG-HSLIANWTPKNYVPHPERQSTLVANTILMNTYSQMAQVSMINT	598
QY	729	SVDRSVCRLGLWGSVSNFFY-HDRDALG-QGYRIYSGGYSGLANSY-FGSSMFLAFTPE	785
DB	599	IAHGAYLFTWGSVSNLFAVHDSCKPIDNWHRSGLYLPGISLHSDHFSCLAAGQ	658
QY	786	VFGSKDYVVCRRNHHACIGSVYLSLTOQALCGSYLFGDAFIRASYGPGNOHMTSY-TFA	844
DB	659	LLGKSDSPITSTE-----TTSVIATVQAQLATPLM-KISAQACYNESIHLELTKYRSFS	712
QY	845	ERSDVRWNNCLAGEICAGLPIVITPSKLYLNELRPVQAEPSVADHESFTESGDQARAF	904
DB	713	KEGFGSHSVAVSGEVCASPIVNSGSLP-SFSFISFKLQGFSGTODGFEESGEIRSF	771
QY	905	KSGHLLNLSVPVGVKPRCSSTHNPKNYSFMAAYICDAYRTISGTETTLTLLSHQETTTDAF	964
DB	772	SASSFRNISLPMGITFEKKQKTENYYFLGAYIQDLKRDVESGPPVLLKNAVSWDAPMA	831
QY	965	HLARHGVVVRGSMVASITSNIEVYGHGREYRDASRGYGLSAGSRVRF	1012
DB	832	NLDSRAYMFLTNQAL-HRLQTLNLVSVYLRGQSHSYSLDLGTTYRF	878
RESULT 35			
ID	Q5L6J6	CHLAB PRELIMINARY; PRT; 840 AA.	
AC	Q5L6J6		
DT	01-FEB-2005	(Tremblrel. 29, Created)	
DT	01-FEB-2005	(Tremblrel. 29, Last sequence update)	
DT	01-FEB-2005	(Tremblrel. 29, Last annotation update)	

DE	Polymorphic outer membrane protein.		
GN	Name=pmpl06; OrderedLocusNames=CAB277;		
OS	Chlamydomphila abortus.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomphila.		
OX	NCBI_TaxID=83555;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RC	STRAIN=S26/3;		
RX	PubMed=15837807; DOI=10.1101/gr.3684805;		
RA	Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,		
RA	Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,		
RA	Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,		
RA	Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.,		
RT	"The Chlamydomphila abortus genome sequence reveals an array of		
RT	variable proteins that contribute to interspecies variation."		
RL	Genome Res. 15:629-640(2005).		
RX	EMBL; CR848038; CAH63728.1; -, Genomic_DNA.		
KW	Complete proteome.		
SK	SEQUENCE 840 AA; 90457 MW; 8FB8CDE411AE1493 CRC64;		
QY	37	YDGETLVSPFVTVIGDPSGTVFVSAGELTKNLNDSIAALPLSCFNLGSLFTVLRGH	96
DB	32	PDGSGTGTQFTAKQSTQAGGTTYNLTADVIIHQVKSTOPA-NTSCFKNSTGDIITFGANH	90
QY	97	SLTFENITSTNGAALSNSAADGLFTIEGPKELSPNSCNLLAVLPAATNKGSTPTTT	156
DB	91	SLIFEDIVSTAGAAISTNTDKTITMSGFNLSP-----IAAPQATT-----	133
QY	157	STPSNGTIYSKTDLLLNNEKFSFYNLSVSGDGAIDAKSLTVQGISKLCVFOENTAQAD	216
DB	134	-----	133
QY	217	GGACQVTVSFSAMANEAFIAFVANVAGVGRGGIAAVQDQGVSSSTSTEDPVPVSPRNT	276
DB	134	-----	133
QY	277	AVEFGNVARVGGGIYSYGNVAFPLNGKTLPLNNVASPVYIAAKQTSQASNTSNYGD	336
DB	134	-----GNA-----IYGIASITIKENNLVP-----DTHNSTAA	162
QY	337	GGAIPE-KNGAAGSNNSGVSFDGEGVVPFSSNVAAGKGAIYAKKLSVANCGPVPFLR	395
DB	163	GGAIHCLTKGATA-----TTLLEKNASWIFRNSSATGGAIHTDLVLTAGGYTLFEN	217
QY	396	NIAND-GGAIYLGESGELSADYGDIIIPDGNLKTAKENAADVNGVTVSSQAIKMGSGG	454
DB	218	NHATQGGAIISAGSGLSLSADEGSIIFRGNTYTDAGNRV-----NNAIYVANG	268
QY	455	KITTLBAKAGHOILFNDPI-EMANGNQAQSSKLLKINDGEG---YTGDIVPANGS---	507
DB	269	KPTKLEAKAQSLFYDPIILVEGAADN-----LEINKANGATVYTGSIIFSGRHH	320
QY	508	-----STLYQNVITEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDPVTPQP	557
DB	321	SPHKMKHVSKEFTQPLTLAGSLILEKGAHLEAKSLTQAGSRILLDTQSS-----	371
QY	558	QQPAAANQLITLSNLHLSLLANNAVTPNPTNPPAQDHPAVIGSTTASGVTISGPI-	616
DB	372	---IKVQENVDIKDLWLSDLDFEPTA-----TH---IASSGDHNVITITGLG	414
QY	617	PFEDLDDTAYRDYDWLGSNQKINVLKQLGTKPPANAPSDLTIGN--EMPK-----	665
DB	415	IPAD-QETPYNHA-LAHNVDBELLQ-----ADKDIKISLVDVPESVRENKDAH	463
QY	666	YGVQSGKWLAD--PNTANN-----GPYTLKATWTKTYNTP--GPERVAS-LVPNSLWGS	715
DB	464	RGYQGSWTIDWKTVPGSTSGGVTLGTATVHWRPTGYIPFGGAQEIITPLVNTLWGN	523
QY	716	ILDIRSAHSAIQASVDGRSYCRGLWGSVSNFPFVHQRDALGQGYRYISGGYSGLANSY-F	774


```

Db 524 FSDINLERTIESLATNSLCSGFWATGKFLYTSQGEKFPVQHNNSGVAIGINKHTL 583
Qy 775 GSSMFLAFTVFGSKSDYVYVCRNHHACISGVY---LSTQALCGSYLFGDA----- 824
Db 584 SENFSAAFSQLFGKDRDQAQVEHQTLSSGLYAHVGTIPML--REFCGDSKQCVPEL 641
Qy 825 -----PIRASVPGNQHMKTSYTFABESDVR-----WDNCLAGEAGLPIVITPSK 872
Db 642 QASPCIPVIFNAQLSYSHS--NNLTIAHEDQTKTGNWSYVATLGSTFVYTLKCP 699
Qy 873 LYLNELRFVQAEPSYADHESFTBEGDQARAFKSHLLNLVSVGVKFDRCSSTHPNKYS 932
Db 700 SILKNSPFIKLVGYSSQRAFTBEGRLRCFSSYLANLALPVGIIKQIGCPKLLAYD 759
Qy 933 FMAAYICDAYRTISGTETT--LLSHQETWTTTDAFLARHGVVVRGSMVASLTSNIEVYGHG 991
Db 760 LSAMTVHDVFRINPETMTFLIGRLAPWTTTATHLDNKALVVGSGRFAVRPNIEVPASG 819
Qy 992 RYERDASRGYGLSAGSVRF 1012
Db 820 NGELRSSHSYNYDFGAKIHF 840

```

RESULT 36

```

Q84FV4 CHLTR
ID Q84FV4 CHLTR PRELIMINARY; PRT; 846 AA.
AC Q84FV4;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Polymorphic membrane protein I (fragment).
GN Name=pmpl;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22483673; PubMed=12595433;
RX DOI=10.1128/IAI.71.3.1200-1208.2003;
RA Stothard D.R., Toth G.A., Batteiger B.B.;
RT "Polymorphic membrane protein H has evolved in parallel with the three
RT disease-causing groups of Chlamydia trachomatis."
RL Infect. Immun. 71:1200-1208 (2003).
DR EMBL; AY104171; AAC029988.1; -; Genomic DNA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR InterPro; IPR005546; Auto transportbeta.
DR InterPro; IPR011427; ChlamPMP M.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF07548; ChlamPMP M; 1.
DR Pfam; PF02415; Chlam PMP; 6.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
FT NON_TER 1
FT NON_TER 846
FT NON_TER 846
SQ SEQUENCE 846 AA; 92007 MW; B0E2F7ED26A75ED CRC64;

```

```

Query Match 14.6%; Score 771.5; DB 2; Length 846;
Best Local Similarity 26.4%; Pred. No. 5.5e-37;
Matches 264; Conservative 145; Mismatches 388; Indels 203; Gaps 33;
Qy 53 DPGTTFVPSAGELTKLNDLSIALPLSCFQNL-----GSTVLGRGHSITYPE 101
Db 2 DPLGETA-----LTKPNHVCTFFEDCTMESLPALCAHASQDDPLVYLGNSYCFVVS 56
Qy 102 NIRTSTGAALSNAADGLFTIEGPKELSPNCNLLAVLPAATTNKGSTPTTTSTPSN 161
Db 57 KLHTDPKEALFKKGD--LSIQNFRFLSFTDCSS-----KESSFS- 95
Qy 162 GTIYSKT--DLILLANNEKFSFYSNIVSGDGGDAIDAKSLTVQGISKLCVFPQNTAQDGGAC 220
Db 96 -IIHQKQGLSLRNNGSMFCRNEAEGSGGAIADAFSLQHNYLPTAFEENSKGNGAI 154

```

```

Qy 221 QVVTSPSAMANEAPAFVAVAGVGGIAAVQDQGGQSSSTSTEDPVVFSRNTAVEF 280
Db 155 QAQT-FSLSRNVSPISFARNADLNGGAICC----- 184
Qy 281 DGNVARVGGIGYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNTGDDGAI 340
Db 185 -----SNLISGNNV-----NPLFF-----TCNSATNGAI 209
Qy 341 FCKNQAQAGNNGSGSVFDGCVVFFSSNVAAGKGGAIYAKKLSVANCGVPQVLFURNIAND 400
Db 210 CCI--SDLNTSEKGLSLACNQETLPASNGAKKGGAIYAKHMVLYNGVPVFINNSAKI 267
Qy 401 GGAIYLGESGELSADYGDIIIPDNLKRTAKENADVNGVTVSSQAISSMGSGKITILR 460
Db 268 GGAIATQGGSLSLAGEBSVLFPQNSQRTSDQLVR-NAIYLEKDAI-----LSSLE 319
Qy 461 AKAGHQILFNDPT-EMANGNPPAQSSKLLKINDGEGYTG-----DIVFA--- 504
Db 320 ARNG-DILFDPVQESSKESPLPSLQASVSTPTATASPLVIQTSANRSVIFSSERL 378
Qy 505 -----NGSSTLYQNVITIEQRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFVTPQ 555
Db 379 SEBEKTPDNLTSQ-LQPIELKSGRLVKDRAVLSPSLSDQPQALLIMEAGTS----- 431
Qy 556 PPOQPPAANQILITLSNLHLSLSLLANNVTPNPPPAQDSHPAVTGSTAGSVTISGP 615
Db 432 -----LKTSSDLKATLSIPLHSL-----DTEKSVTHIAP 461
Qy 616 -----IFFEDL-DDTAYRDYDWMGSKINVLKQLGKTKPPANAPSDLTIGNEMPKYGY 668
Db 462 NLSIQKIFLNSGDNFENVLLSKQN---NIPLLTLPKQSHLHLPDGNLSSHPGY 517
Qy 669 QGSKLAWDPNTANGPYTLKATWTKTYNPGPERVASLVPNSLWGLSILDIRSAHSAIQ 728
Db 518 QGDWTFSW--KDSDEG-HSLIANWTPKVVYPHPERQSTLVANTLWNTYSDMQAVQSMINT 574
Qy 729 SVDGRSYCRGLWVGVSNNPFY-HDRDALG-QGTRYISGGYSLGAN8Y-FCSSMPLGAPTE 785
Db 575 TAHGGAYLFGTWGSAVSNLFYAHDSGKPIDNWHHRSGLYLPGISLTHSLDDHDFCLAAQ 634
Qy 786 VFGRSKDYVYCRNHHACISGVYLSLTOALCGSYLFGDAFIRASYPGNQHMKTSY-TFA 844
Db 635 LLGKSSDPTTSTB-----TTSYIATVQAQATSLM-KISAQACYNESIHELKTKYRSP 688
Qy 845 EESDVRMDNCLAGEIGAGLPVITPSKLYNLRLPFVQAEPSYADHESFTBEGDQARAF 904
Db 689 KEGFGSMHSAVSGEVCALIPVNSGSLF-88FSIFSKLQGFSGTQDGPESSEGEIRSF 747
Qy 905 KSGHLLNLVSVGVKFDRCSSTHPNKYSFMAAYICDAYRTISGTETTLLSHQETWTTDAP 964
Db 748 SASSFRNISLPIGITFEKKSQKTRNYFFLGAVIQDLKRDVSGPVVLLKNAVSWDAPNA 807
Qy 965 HLAHGVVVRGSMYASLTNIEVYGHGRYBYRDAISRGYGL 1004
Db 808 NLASRAYMFLTNQRAL-HRLQTLNLNVSVVLRQGSYSYL 846

```

RESULT 37

```

PMPI CHLTR
ID PMPI CHLTR STANDARD; PRT; 878 AA.
AC O84882;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable outer membrane protein pmpl precursor (Polymorphic membrane
DE protein 1).
GN Name=pmpl; OrderedLocusNames=CT874;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=D/UM-3/Cx;

```

RA MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.B., Zhao Q., Koonin S.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RL Chlamydia trachomatis";
RL Science 282:754-759(1998).
CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (potential).
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL: AB001361; AAC68472.1; -; Genomic_DNA.
CC PIR: B71460; B71460.
CC InterPro: IPR005546; Auto transptbeta.
CC InterPro: IPR011427; ChlamPMP_M.
CC InterPro: IPR003368; Chlamydia_PMP.
CC Pfam: PF03797; Autotransporter; 1.
CC Pfam: PF02415; Chlam_PMP; 6.
CC Pfam: PF07548; ChlamPMP_M; 1.
CC TIGRfam: TIGR01376; POMP_repeat; 6.
KW Complete proteome; Membrane; Multigene family; Outer membrane; Signal.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 878 Probable outer membrane protein pmp1.
FT SEQUENCE 878 AA; 95593 MW; DF1F1A31707EE48B CRC64;
SQ

Query Match 14.8%; Score 771.5; DB 1; Length 878;
Best Local Similarity 26.5%; Pred. No. 5.8e-37;
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

QY 53 DPSTGVFSAGELTLKLNLSIAALPLSCFGLNLT-----GSFTVLGRGHSLTPE 101
DB 26 DPLGETA-----LTKNPNHVCTTFEDCTMESLFPALCNHASQDDPLVLGNSYCFVS 80

QY 102 NIRTSTNGAALNSAADGLFTIEGFKELSPNCNLSLLAVLPAATTNKGSGQTPTTSTPSN 161
DB 81 KLHITDPKEALPKBKGD--LSIQNFRFLSFTDCSS-----KESSPS- 119

QY 162 GTIYSKT-DLLILNNKFPYSNLSVGGDAIDAKSLTWGISKLVCFQENTAAQAGGAC 220
DB 120 -IIHQKNGQLSLNNGSMSPCRNHAEGSGGAIADAFSLQHNYLFTAFENSSKGGGAI 178

QY 221 QVVTFSAMANEAPIAFVANVAGVGGIAAVQDGGQVSSSTSTEDPVVFSRNTAVF 280
DB 179 QAQT-FSLSRNVSPISFARNRDLNGAICC----- 208

QY 281 DGNVAVGGGIYGYGNVAFNLNGKTLFLNNVAFSPVIAAKQPTSGQASNTSNNGYDGGAI 340
DB 209 -----SNLISGNNV-----NLPFP-----TGNATNGGAI 233

QY 341 FCKNGAQAAGNSGVSFDEGVVFPSSNVAAGKGGAIYAKLSVANCGVQFLRIAND 400
DB 234 CCI--SDLNTSEKGLSLACNQETLFPASNSAKGKGGAIYAKHVLRYNGFVSPINNSAKI 291

QY 401 GGAIYLGESGELSLADYGDIIIDGNLKRKTAKENADVNGVTVSSQASNGSGGKITTLR 460
DB 292 GGAIATQSGGSLSLAGEGSLVFQNNNQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343

QY 461 AKAGHQILFNDPI-EMANGNNQAQSSKLKINDGEGVTG-----DIVFA----- 504
DB 344 ARNG-DILFDPPIVOESSEKESPLPSSLSQASVTSPTATAPLVIQTSANRSVIFSSERL 402

QY 505 -----NGSSTLYQNVTTIEQGRIVLREKAKLSVNSLSQTGGS-LYMEAGSTWDFTPQ 555
DB 403 SEBEKTPDNLTSQLOQPIELKSGRLVKDRAVLSPSLSDQPOALLIMEAGTS----- 455

QY 556 PPOQPPAANQLITLSNLHLSLSLLANNVNTNPTNPQADSHPAVIGSTTAGSVTISGP 615

DB 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHIAP 485
QY 616 -----IPEDEL-DDTAYDRYDMLGSKNQKINVLKQLGKTGPPANAPSDLTL--GNEMPKY 666
DB 486 NLSQKIFLSNGDENFYENVVLLSKEQN-NIPLLT-----SKEQSHLLPLDGNLSSH 539
QY 667 GYQGSWKLAWDPNTANNPYTLKATWTKTGYNPGRVASLVPNSLWGSILDIRSAHSAI 726
DB 540 GYQGDWTFPSW--KDSDEG-HSLIANWTPKNTVPHPERQSTLVANTLWNTYSDMQAVQSMI 596
QY 727 QASVDGRYCGRLWVGVSNEFY-HDRDALG-QGVRYISGGVSLGANYS-FGSSNMFGLAF 783
DB 597 NTIAHGGAFLGTGWSAVSNLFYHDSGKPIDNWHHSLGLVLFQISTHSDHDFCLAA 656
QY 784 TEVFGSRKDYVVCNHHACIGSVVLSYQALCGSYLFGDAFIRASYGFGNQHMKTSY-T 842
DB 657 GQLLGKSSDSFITSTE-----TTSYIATVQAQLATPLM-KLSAQACYNESIHELKTKYS 710
QY 843 FAESDVRWNNCLAGEIGAGLPVITPSKLYNELRPFVQAERFVADHESFTEGGDQAR 902
DB 711 FSKGFGGWSHVSVAVSGEVCASPIVNSGGLP-SSFSIFSCLKQGFSGTQDGFEESSGEIR 769
QY 903 AFKSGHLLNLSVPVGVKEDRCSSHTPNKYSFMAAYICDAYRTISCTETTLTLLSHQETWTD 962
DB 770 SFSASSFRNISLPMGITFEKKSQKTRNYYPFLGAYIQDLKRDVSGPVLVLLKNVSWDAP 829
QY 963 AFHLARHGVRGSMYASILTSNIEVYGHGRYERDASRGYGLSAGSRVRF 1012
DB 830 MANLDSRAYMFLTNQRAL-HRLQTLNLSVYLRQSHSYSLDLGTTYRF 878

RESULT 38
Q4VWS2_CHLTR
ID Q4VWS2_CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWS2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymorphic membrane protein 1.
GN Name=pmp1;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
ON NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=UW-12, UW-202, UW-31, UW-4, and UW-3;
RA Gomes J.P., Borrego M., Dean D.;
RT "Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY299433; AAQ74449.1; -; Genomic DNA.
DR EMBL: AY299434; AAQ74450.1; -; Genomic DNA.
DR EMBL: AY299436; AAQ74452.1; -; Genomic DNA.
DR EMBL: AY299431; AAQ74447.1; -; Genomic DNA.
DR EMBL: AY299445; AAQ74461.1; -; Genomic DNA.
DR InterPro: IPR005546; Auto transptbeta.
DR InterPro: IPR011427; ChlamPMP_M.
DR InterPro: IPR003368; Chlamydia_PMP.
DR Pfam: PF03797; Autotransporter; 1.
DR Pfam: PF02415; Chlam_PMP; 6.
DR TIGRfam: TIGR01376; POMP_repeat; 6.
SQ SEQUENCE 878 AA; 95593 MW; DF1F1A31707EE48B CRC64;

Query Match 14.8%; Score 771.5; DB 2; Length 878;
Best Local Similarity 26.5%; Pred. No. 5.8e-37;
Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

QY 53 DPSTGVFSAGELTLKLNLSIAALPLSCFGLNLT-----GSFTVLGRGHSLTPE 101
DB 26 DPLGETA-----LTKNPNHVCTTFEDCTMESLFPALCNHASQDDPLVLGNSYCFVS 80

QY 102 NIRTSTNGAALNSAADGLFTIEGFKELSPNCNLSLLAVLPAATTNKGSGQTPTTSTPSN 161

Db 81 KLHITDPKEALFKEGD--LSIQNFRPLSFPTDCSS-----KESSPS- 119
Qy 162 GTIYSKT-DLLLNNEKSFYSNLVSGDGGAIADAKSLTVQGISKLCVFOENTAOADGAC 220
Db 120 -IIHQKQGLSLNRNGSMFSCRNHAEGSGGAIADAFSLQHNLYLFTAFENSSKNGGAI 178
Qy 221 QVVTSPSAMNAPAFIAFVANVAGVGGGIAAVQDQCGQVSSSTSTEDPVVFSRNTAVEF 280
Db 179 QAQT-FSLSRNVSPISFARNRADLNGAICC----- 208
Qy 281 DGNVARGGIGYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNYGDGAI 340
Db 209 -----SNLICSGNV-----NPLPF-----TGSATNGGAI 233
Qy 341 FCNKAQAGSNNGSVSFDGEGVVPSSNVAAGKGAIYAKKLVSANGCPVQFQFLNNTAND 400
Db 234 CCI--SDLTSEKGLSLACNQETLFASSAKGKGAIYAKHMYLRYNGPVSFINNNAKI 291
Qy 401 GGAIYLGESGELSLADYDIIFDGNLKRKTAKENAADVNGVTVSSQAISSMGSGKITTLLR 460
Db 292 GGAIYQSGGSLILAGEGSLVFNQNSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343
Qy 461 AKAGHQILFNDDPI-EMANGNNOQAQSSKLLKINDGEGYT-----DIVFA----- 504
Db 344 ARNG-DILFPDPIVQESSKESPLSSLOASVTSPTATASPLVIQTSANRSVIFSSERL 402
Qy 505 -----NGSSTLYQNTYIEGRIVLREKAKLSVNSLSQSGS-LYNEAGSTWDFVTPQ 555
Db 403 SEEEKTPDNLTSQLQPIELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS----- 455
Qy 556 PPQOPPAANQLITLSNLHLSLLANNVTPNPPAQAQSHPAVIGTSITAGSVTISGP 615
Db 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485
Qy 616 -----IPFDL-DDTAYDRYDMGLSNQKINVLKQLGKTPPANAPSDLTL--GNEMPKY 666
Db 486 NLSIQIKIFLSNGDENFENVVLLSKEQN-NIPLTL-----SKEQSHLHLPDGNLSHP 539
Qy 667 GYQGSWKLAWDPNTANNPYTLKATWTKYNGPGRVAVSLVPSNLSGSLIDIRSAHAI 726
Db 540 GYQGDWTFW--KDSDEG-HSLIANWTPKNVPHPERQSTLVANTLNTYSDMQAVQSMI 596
Qy 727 QASVDGRSYCRGLWVSGVSNFFY-HDRDALG-QGYRYISGYSGLANSY-FGSSMFGAP 783
Db 597 NTIAHGGAFLFTWGSVAVSNLFYAHDSGKPIDNWHRSGLYFGISTHSDHSPCLAA 656
Qy 784 TEVFGSKDYVVCNHNHACIGSVYLSLTOQALCGSYLFGDAPIRASYFGNGHMKTSY-T 842
Db 657 QOLLGKSSDSFITSTE-----TTSYIATVQALATPLM-KISAQACYNESITHELTKYRS 710
Qy 843 PABSDVWDDNCLAGEIGAGLPVITPSPKLYLNLPRFPVQAFSYADHESFTTEGDOAR 902
Db 711 FSKEGFGSWHSVAVSGEVCASIPVNSGSLF-SFSFISFKLQFGSGTQDGFESSGDIR 769
Qy 903 AFKSHLLNLVSPVGVKEDRCSSHPNKYSFMAAYICDAYTISGETTETLLSHOETWTD 962
Db 770 SFSASSFNISLPMGITEKSKQKTRNYTFFLGAAYIQDLKRDVESGPPVLLKNAVSWDAP 829
Qy 963 APHLARHGVVVRGMSVYASLTENIEVYGHREYDASRGYGLSAGSRVF 1012
Db 830 MANLDSRAYMFLTNQORAL-HRLQTLNLVSVYLRQGSYSYSLDGLGTTTF 878

RESULT 39

Q4VWR1.CHLTR PRELIMINARY; PRT; 878 AA.
AC Q4VWR1.
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Polymorphic membrane protein 1.
GN Nene=omp1;
OS Chlamydia trachomatis.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=UW-57;

RA Gomes J.P., Borrego M., Dean D.;

RT *Phylogenetic analysis of pmp's genes of Chlamydia trachomatis.*;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY299444; AAQ74460.1; -; Genomic DNA.

SQ SEQUENCE 878 AA; 95534 MW; DF1FIA333B13501B CRC64;

Query Match 14.6%; Score 770.5; DB 2; Length 878;

Best Local Similarity 26.5%; Pred. No. 6.6e-37;

Matches 268; Conservative 145; Mismatches 390; Indels 207; Gaps 35;

Qy 53 DPSGTTVFSAGELTKNLDSIAALPLSCPNLL-----GSFTVLGRGSLTFE 101

Db 26 DPLGETA-----LTKPNHVCTTFEDCTMESLFPALCAHASQDDPLYVLGNSYCMFVS 80

Qy 102 NIRTSTGAALNSAADGLFTIEGFKLSFNSCNLSLAVLPAAATNKGSTPTTTSTPSN 161

Db 81 KLHITDPKEALFKEGD--LSIQNFRPLSFPTDCSS-----KESSPS- 119

Qy 162 GTIYSKT-DLLLNNEKSFYSNLVSGDGGAIADAKSLTVQGISKLCVFOENTAOADGAC 220

Db 120 -IIHQKQGLSLNRNGSMFSCRNHAEGSGGAIADAFSLQHNLYLFTAFENSSKNGGAI 178

Qy 221 QVVTSPSAMNAPAFIAFVANVAGVGGGIAAVQDQCGQVSSSTSTEDPVVFSRNTAVEF 280

Db 179 QAQT-FSLSRNVSPISFARNRADLNGAICC----- 208

Qy 281 DGNVARGGIGYSYGNVAFNNGKTLFLNNVASPVYIAAKOPTSQASNTSNNYGDGAI 340

Db 209 -----SNLICSGNV-----NPLPF-----TGSATNGGAI 233

Qy 341 FCNKAQAGSNNGSVSFDGEGVVPSSNVAAGKGAIYAKKLVSANGCPVQFQFLNNTAND 400

Db 234 CCI--SDLTSEKGLSLACNQETLFASSAKGKGAIYAKHMYLRYNGPVSFINNNAKI 291

Qy 401 GGAIYLGESGELSLADYDIIFDGNLKRKTAKENAADVNGVTVSSQAISSMGSGKITTLLR 460

Db 292 GGAIYQSGGSLILAGEGSLVFNQNSQRTSDQGLVR-NAIYLEKDAI-----LSSLE 343

Qy 461 AKAGHQILFNDDPI-EMANGNNOQAQSSKLLKINDGEGYT-----DIVFA----- 504

Db 344 ARNG-DILFPDPIVQESSKESPLSSLOASVTSPTATASPLVIQTSANRSVIFSSERL 402

Qy 505 -----NGSSTLYQNTYIEGRIVLREKAKLSVNSLSQSGS-LYNEAGSTWDFVTPQ 555

Db 403 SEEEKTPDNLTSQLQPIELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS----- 455

Qy 556 PPQOPPAANQLITLSNLHLSLLANNVTPNPPAQAQSHPAVIGTSITAGSVTISGP 615

Db 456 -----LKTSSDLKLATLSIPLHSL-----DTEKSVTHAP 485

Qy 616 -----IPFDL-DDTAYDRYDMGLSNQKINVLKQLGKTPPANAPSDLTL--GNEMPKY 666

Db 486 NLSIQIKIFLSNGDENFENVVLLSKEQN-NIPLTL-----SKEQSHLHLPDGNLSHP 539

Qy 667 GYQGSWKLAWDPNTANNPYTLKATWTKYNGPGRVAVSLVPSNLSGSLIDIRSAHAI 726

Db 540 GYQGDWTFW--KDSDEG-HSLIANWTPKNVPHPERQSTLVANTLNTYSDMQAVQSMI 596

Qy 727 QASVDGRSYCRGLWVSGVSNFFY-HDRDALG-QGYRYISGYSGLANSY-FGSSMFGAP 783

Db 597 NTIAHGGAFLFTWGSVAVSNLFYAHDSGKPIDNWHRSGLYFGISTHSDHSPCLAA 656

Qy 784 TEVFGSKDYVVCNHNHACIGSVYLSLTOQALCGSYLFGDAPIRASYFGNGHMKTSY-T 842

Db 657 QOLLGKSSDSFITSTE-----TTSYIATVQALATPLM-KISAQACYNESITHELTKYRS 710

Qy 843 PABSDVWDDNCLAGEIGAGLPVITPSPKLYLNLPRFPVQAFSYADHESFTTEGDOAR 902

Db 711 FSKEGFGSWHSVAVSGEVCASIPVNSGSLF-SFSFISFKLQFGSGTQDGFESSGDIR 769

Qy 903 AFKSHLLNLVSPVGVKEDRCSSHPNKYSFMAAYICDAYTISGETTETLLSHOETWTD 962

Db 770 SFSASSFNISLPMGITEKSKQKTRNYTFFLGAAYIQDLKRDVESGPPVLLKNAVSWDAP 829

Qy 963 APHLARHGVVVRGMSVYASLTENIEVYGHREYDASRGYGLSAGSRVF 1012

Db 830 MANLDSRAYMFLTNQORAL-HRLQTLNLVSVYLRQGSYSYSLDGLGTTTF 878

Db	711	PSKEGFGSHSVASGEVCASIPVYNSGSLF--SSFSIFSKLQFGSQTDGPFSSGBIR	769
Qy	903	AFKSGHLNLNLSVPVGVKDFRCSTTHPNKYSFMAAICDAYRTISGTETTLTSHQETWTD	962
Db	770	SFSAGSFRNISLPMGITTEKSKQTRNYYIFLGAYIQDLKRDVSGPVVLLKNVSWGAP	829
Qy	963	AFLHARHGTVVGRGVSAYSTNIEVVGHRVYRDASRGYGLSAGSRVRF	1012
Db	830	MANLDSRAYMFLTNQRAL--HRLQTLNLSVYVLRGQSHSYSLDLGTTTFP	878
RESULT 40			
Q84FU2	CHLTR		
ID	Q84FU2	CHLTR PRELIMINARY; PRT; 846 AA.	
AC	Q84FU2;		
DT	01-JUN-2003	(TrEMBLrel. 24, Created)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Polymorphic membrane protein I (Fragment).		
GN	Name=pmpl;		
OS	Chlamydia trachomatis.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=813;		
BN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22483673; PubMed=12595433;		
RX	DOI=10.1128/JAI.71.3.1200-1208.2003;		
RA	Stoithard D.R., Toth G.A., Betteiger B.E.;		
RT	"Polymorphic membrane protein H has evolved in parallel with the three		
RT	disease-causing groups of Chlamydia trachomatis.";		
RL	Infect. Immun. 71:1200-1208(2003).		
DR	EMBL; AY184173; AAC029990.1; -; Genomic DNA.		
DR	GO; GO:0019867; C:outer membrane; IEA.		
DR	InterPro; IPR005546; Auto transptbeta.		
DR	InterPro; IPR011427; ChlamPMP M.		
DR	InterPro; IPR003368; Chlamydia_PMP.		
DR	Pfam; PF03797; Autotransporter; 1.		
DR	Pfam; PF07548; ChlamPMP M; 1.		
DR	Pfam; PF02415; Chlam_PMP; 6.		
DR	TIGRfam; TIGR01376; POMP_repeat; 6.		
FT	NON_TER	1	
FT	NON_TER	846	
FT	NON_TER	846	
SQ	SEQUENCE	846 AA; 92023 MW; OA50AF8E507E8PF6 CRC64;	
Query Match 14.6%; Score 769.5; DB 2; Length 846;			
Best Local Similarity 26.4%; Pred. No. 7.2e-37;			
Matches 264; Conservative 145; Mismatches 388; Indels 203; Gaps 33;			
Qy	53	DPGTTVFSAGELTKVLDNSIAALPLSCFGLI-----GSFTVLGRGHSITPE	101
Db	2	DPLGETA-----LITKPNHVCTTFEDCTMESLPALCAHSAQDDPLYLGNISYCFVS	56
Qy	102	NIRTSTNGAALSAAAGLFTIEGFKLSFNCNLSLAVLPAAATNKGSTPTTITSPSN	161
Db	57	KLHTDPKEALFKEKGD--LSIQNFRFLSFDCCS-----KESFPS-95	
Qy	162	GTIYSKT-DLLLNNEKFSYSLNLSVGDGAIDAKSLTVQGISKLVPQENTAOAGGAC	220
Db	96	-LIHQKQGLSLRNNGSNFCENHAEGSGGAI SADAFSLQHNLYLFTAFENSSKNGGAI	154
Qy	221	QVVTFSAMANEAFIAFVANVAVRGGIAAQVQGGQVSSSTEDPVPVPSRTAVEP	280
Db	155	QAQT-FSLSRNVSPISFSRNRADLNGGAIC-----	184
Qy	281	DGNVAVCGGIYSYGNVAFNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNGDGA	340
Db	185	-----SNLICGNV-----NPLFF-----TGNATNGAI	209
Qy	341	PKNGAAGSNNSGSVSPDGGVVFSSNVAAGKGAIYAKKLVSANGCPVQFLRNIAND	400
Db	210	CCI--SDLNTSEKLSLACHQETLFPASNSAKEKGAIYAKHVLRYNGPVSFINNSAKI	267
Qy	401	CGAIYLGESBELSLADYGDIIIPDGNLKR TAKENAADVNGVTVSSQAISMGSGGKIITLR	460

Db	268	GGATAIQGGSLSLAGEGSLFQWNSQRTSDQGLVR--NATYLEKDAI-----LSSLR	319
Qy	461	AKAGHQILLFNDPI--EMANGNNQPAQSSKLLKINDGEVYG-----DIVFA----	504
Db	320	ARNG--DILLFPPIVOESSKESPLPSSLOASVTSPTPATASPLVITQTSANRSVIFSSERL	378
Qy	505	-----NGSSTLYQNVITIEQGRIVLREKAKLSVNSLSQTGGG--LYMEAGSTWDFVTPQ	555
Db	379	SEEEKTPDNLTSQLOQPTELKSGRLVKDRAVLSAPLSQDPQALLIMEAGTS-----	431
Qy	556	PPQPPAANQLITLSNLHLSLSSLANNAVTPNPPAQDSHPAVIGSTTAGSVTISGP	615
Db	432	-----LKTSSDLKATLSIPLHSL-----DTEKSVTIHAP	461
Qy	616	-----IPEDL--DDTAYDRYDMLGSNOKINVLKQLGTGKPPANAPSDLTLGNEMPKYGY	668
Db	462	NLSIQIKIFLSNGSDENFYENVLLSKEQN-----NIPLITLTPKEQSHLHLPDGNLSHFY	517
Qy	669	QGSWKLAWDPNTANNPGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDIRSAHSAIOA	728
Db	518	QGDWTFW--KDSDEG--HSLIANWTPKNVPHPERQSTLVANTLWNTYSDMQAVQSMINT	574
Qy	729	SVDGRSYCRGLWVSGVSNPFY--HRRDALG--QGYRYISGYSILGANSY--FGSSMFGLAFT	785
Db	575	TAHGGAYLFGTWGSAVSNLFYAHDSGSKFIDNWHRSLSGLYLFGISLTHSLDDHSFCLAAQ	634
Qy	786	VFRSKDYVVCRSNHHACIGSVYLVSTQOALCGSYLFGDAFIRASYFGNGHMKTSY--TFA	844
Db	635	LLGKSDSPITSTE-----TTSYIATVQALATSLM--KLSAQACYNESIHETKTKYRSP	688
Qy	845	EBSDVRWDMNCLAGETGAGLPVITPTSKLYLNLRLPFFVOAEFSYADHESFTEEGDOARAF	904
Db	689	KEGFGSHSVASGEVCALIPVNSGSLF--SSFSIFSKLQFGSQTDGPFSSGBIRSF	747
Qy	905	KSGHLNLNLSVPVGVKDFRCSTTHPNKYSFMAAICDAYRTISGTETTLTSHQETWTDAP	964
Db	748	SASSFRNISLPIGITPEKSKQTRNYYIFLGAYIQDLKRDVSGPVVLLKNVSWDAPWA	807
Qy	965	HLARHGTVVGRGVSAYSTNIEVVGHRVYRDASRGYGL	1004
Db	808	NLASRAYMFLTNQRAL--HRLQTLNLSVYVLRGQSHSYSL	846

Search completed: May 13, 2006, 12:19:49
Job time : 263 secs

THIS PAGE BLANK (USPTO)